

MONTHLY NOTICES
OF THE
ROYAL ASTRONOMICAL SOCIETY

Vol. 98 No. 8

MEETING OF 1938 JUNE 10

Dr. H. Spencer Jones, President, in the Chair.

The election by the Council of the following Fellows was duly confirmed :—

Canon Edward Baunton Ford, Chancellor of the Diocese of Bloemfontein and Rector of Wepener, O.F.S., South Africa ;
Robert R. McMath, Director of the McMath-Hulbert Observatory, Lake Angelus, Michigan, U.S.A. ; and
Mentore Maggini, Director, R. Osservatorio Astronomico, Collurania-Teramo, Italy.

Eighty-three presents were announced as having been received since the last meeting, including :—

Moulton (Forest Ray), "Consider the Heavens" (presented by the author).

TWO SOLAR ERUPTIONS.

J. C. Dobbie, W. Moss and A. D. Thackeray.

A fairly close association between bright chromospheric eruptions and disturbances in the terrestrial ionosphere has been established by a number of observers during the past few years. There is some evidence that the correlation is closest for the most intense solar eruptions,* but since most intensity values hitherto given have been visual estimates it is highly desirable to replace these by accurate photometric measurements. This note describes preliminary results of a programme initiated at the Solar Physics Observatory, Cambridge, in collaboration with the Field Station of the Cavendish Laboratory, with the dual object of (a) photometric measurement of the variation in solar $H\alpha$ radiation during an eruption for comparison with the simultaneous variation in the reflection coefficient of radio-echoes from the terrestrial ionosphere, (b) detection, if possible, of an increase in the continuous light at the time of an eruption, particularly in the ultra-violet.

(a) Technique has been developed whereby, as soon as an eruption is announced by the spectrohelioscope observer, a succession of spectroheliograms may be taken of a small area (about $7'$ square) centred on the eruption at intervals of a minute. Eighteen exposures can be obtained on one plate, which is then brush-developed together with another plate from the same box, calibrated in a tube sensitometer. Two calibration exposures of 3 and 5 seconds are given, the source being a 1500-watt lamp; these short exposures mimic to some extent the rise and fall in intensity to which the plate is exposed as it moves past the secondary slit† of the spectroheliograph. The plates used, Ilford Astra VIII, have a comparatively narrow band of sensitivity including $H\alpha$, and the use of a red screen in the sensitometer ensures calibration at the proper wave-length.

(b) The direct photographs have been taken with an 18-foot focus aluminised mirror, figured by Burch, giving a two-inch image. A polished plate of Vita glass, silvered on both sides, was placed in the beam about eight inches in front of the plate; this filter transmitted a narrow spectral region around 3220 Å. about 20 to 50 Å. broad (depending on the thickness of the silver). Sputtered films, prepared with the help of Mr. W. H. Manning, showed a marked superiority, both in freedom from pinholes and in evenness, over chemically deposited silver. Ilford Gaslight Lantern slides were used; these plates have the useful property of being practically insensitive to wave-lengths longer than 4000 Å. Extreme contrast and fine grain render them highly suitable to this work. Exposures are of the order of one second; in general the resolving power is determined entirely by atmospheric conditions.

* H. W. Newton and H. J. Barton, *M.N.*, 97, 603, 1937.

† The slit width corresponded to 0.15 Å. at $H\alpha$.

Eruption of 1938 March 31.—At $10^{\text{h}} 28^{\text{m}}$ a bright eruption of very small area was observed with the spectrohelioscope about $12''$ within the east limb close to the following member of a pair of medium-sized spots. $H\alpha$ spectroheliograms taken at $9^{\text{h}} 18^{\text{m}}$ and $10^{\text{h}} 8^{\text{m}}$ had shown some activity in this region but nothing that could be classified as an eruption. Direct photographs had been obtained at approximately ten-minute intervals from $9^{\text{h}} 47^{\text{m}}$ onwards, which showed several bright faculae in the vicinity, but at a

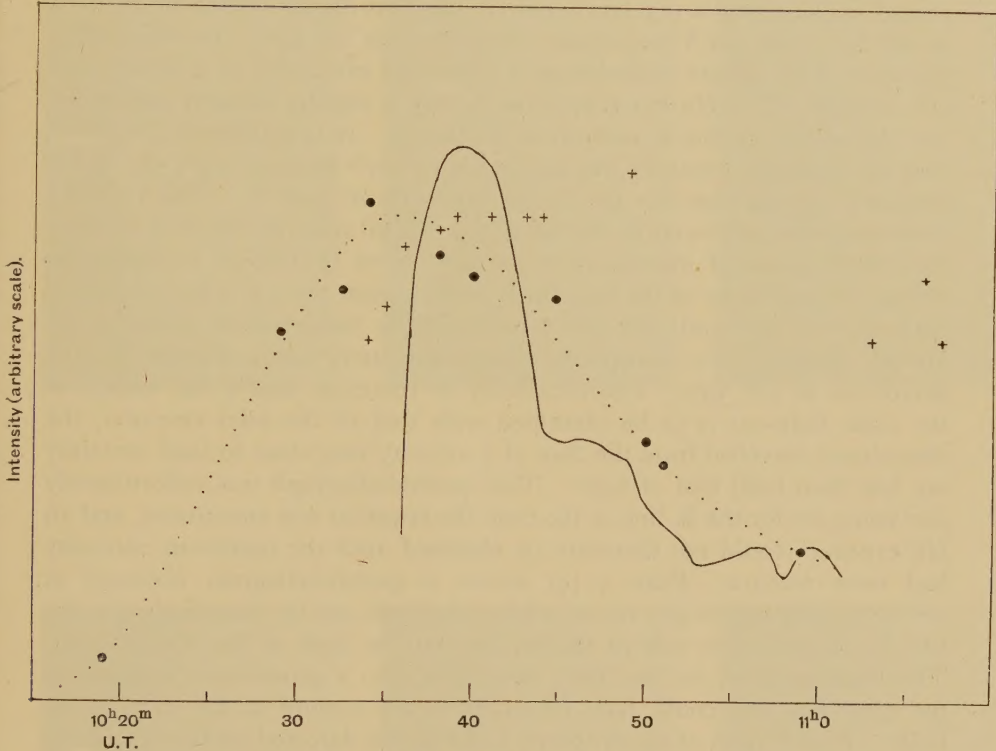


FIG. 1.—Eruption of 1938 March 31.

—————	Ionospheric absorption (5 megacycles).	
+ +	$H\alpha$ eruption intensity	} Visual Estimates.
• •	u.v. „ „	

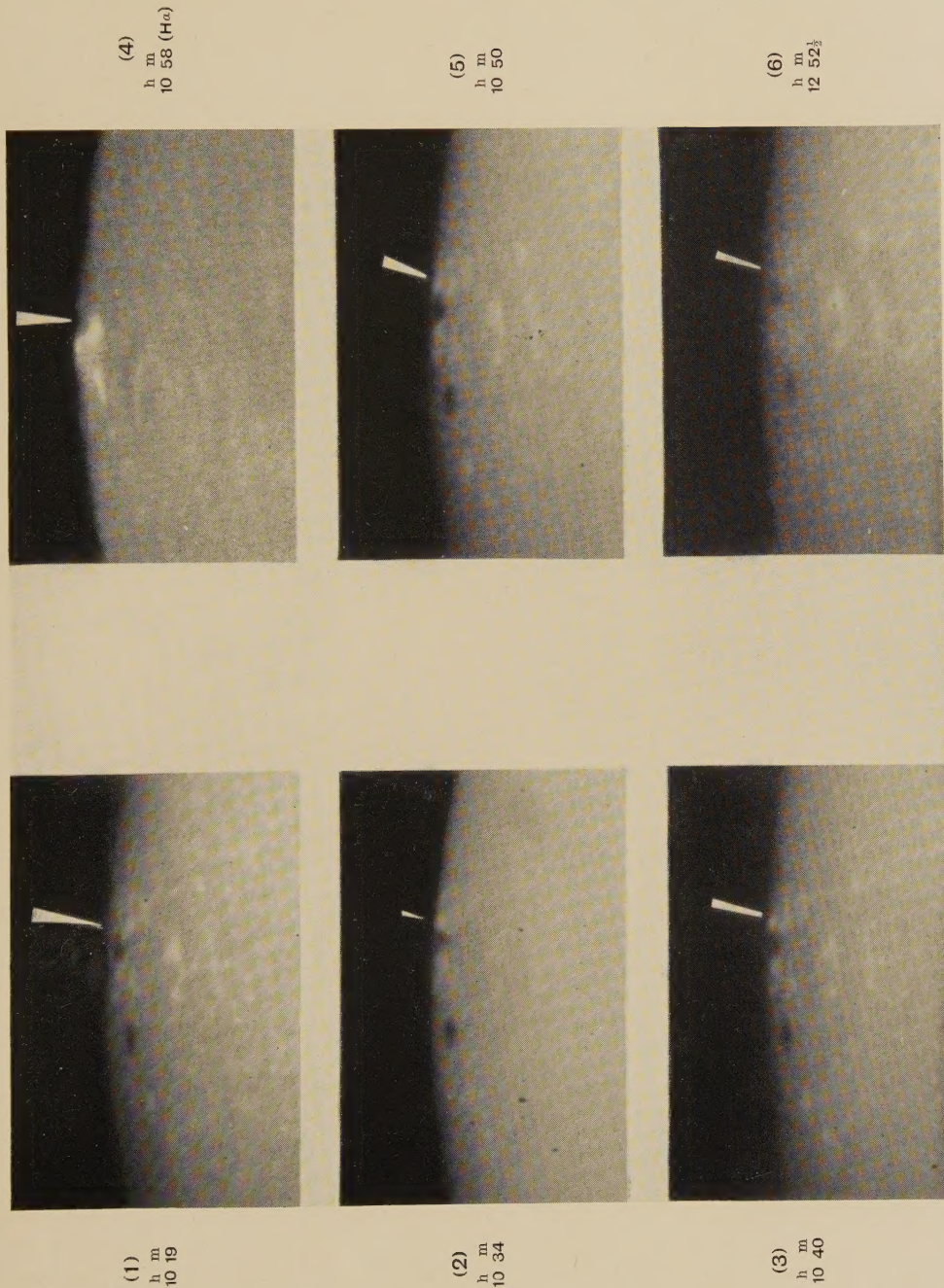
telephonic signal from the spectrohelioscope observer the interval between exposures was shortened to two to three minutes. This series of exposures showed a bright facula coinciding in position with the brightest point of the eruption. It was invisible on the earliest exposures, just detectable at $10^{\text{h}} 19^{\text{m}}$, and in general followed roughly the intensity variation in $H\alpha$ (Plate 5). No calibration was available at the time; (this can now be obtained with the aid of a quartz mercury arc and tube sensitometer). Fig. 1 shows the intensity variation in this temporary facula and in $H\alpha$, both based on visual estimates, the former by inspection of the u.v. plates and comparison with neighbouring faculae, the latter being recorded at the spectrohelioscope by means of a stepped wedge held in front of the eye-piece and comparison

with the neighbouring disc.* The figure also shows the absorption curve of the radio-echoes on 5 megacycles kindly supplied by Professor E. V. Appleton and Mr. W. R. Piggott of the Field Station of the Cavendish Laboratory in advance of publication. From these curves it might be imagined that the solar eruption began some twenty minutes before the ionospheric disturbance. But the time at which a fade-out becomes perceptible depends largely on the frequency used; in this instance, the record on long-waves (15 km.) showed the disturbance beginning as early as $10^h 15^m$, that on 1.8 megacycles (167 metres) at $10^h 24^m$. For this reason the time of maximum disturbance is somewhat preferable as a criterion of coincidence. The $H\alpha$ curve appears to have a slightly delayed maximum, but the reality of this is somewhat doubtful.† It is significant, however, that the hydrogen intensity was as great at $11^h 07^m$ as at $10^h 34^m$, which was certainly not the case for the facula intensity at 3220 Å. The recovery after maximum is slowest in $H\alpha$ light, and is in all cases slower than the rise. The exact epoch of maximum in the u.v. curve is difficult to determine owing to the density of the exposures, but one can place it with confidence between $10^h 34^m$ and $10^h 40^m$ (confirmed by independent estimates by Dr. H. Brück); the ionospheric absorption curve has a sharply defined maximum at $10^h 39^m$. The conclusion is therefore that if the source of the radio fade-out is to be identified with that of the solar eruption, the disturbance travelled from the Sun at a velocity very close to (and certainly not less than half) that of light. The spectroheliograph was unfortunately just being set for the K line at the time the eruption was announced, and an $H\alpha$ exposure could not therefore be obtained until the maximum intensity had been reached. Plate 5 (4) shows a spectroheliogram obtained at $10^h 58^m$; the two bright nuclei within the limb can be identified with the two faculae on either side of the spot nearer the limb in the u.v. pictures. The third nucleus, on the limb, developed into a prominence connecting the other two and could later be traced to 0.6 minute of arc outside the limb. This portion of the eruption could not be detected on the u.v. plates at any stage. The prominence was first noted at the spectrohelioscope at $11^h 52^m$ as a jet projecting radially from the limb. In the course of a few minutes it developed into an arch which later became double, the two components overlapping, then single again, broke into two halves, re-formed and divided into three overlapping arches. A short-lived prominence in the same region was noted earlier at $10^h 41^m$ when the eruption was near its peak. This prominence showed a velocity of 58 km./sec. towards the Earth.

The very close proximity to the limb renders intensity measures of this eruption extremely difficult, but measurement of a spectroheliogram taken

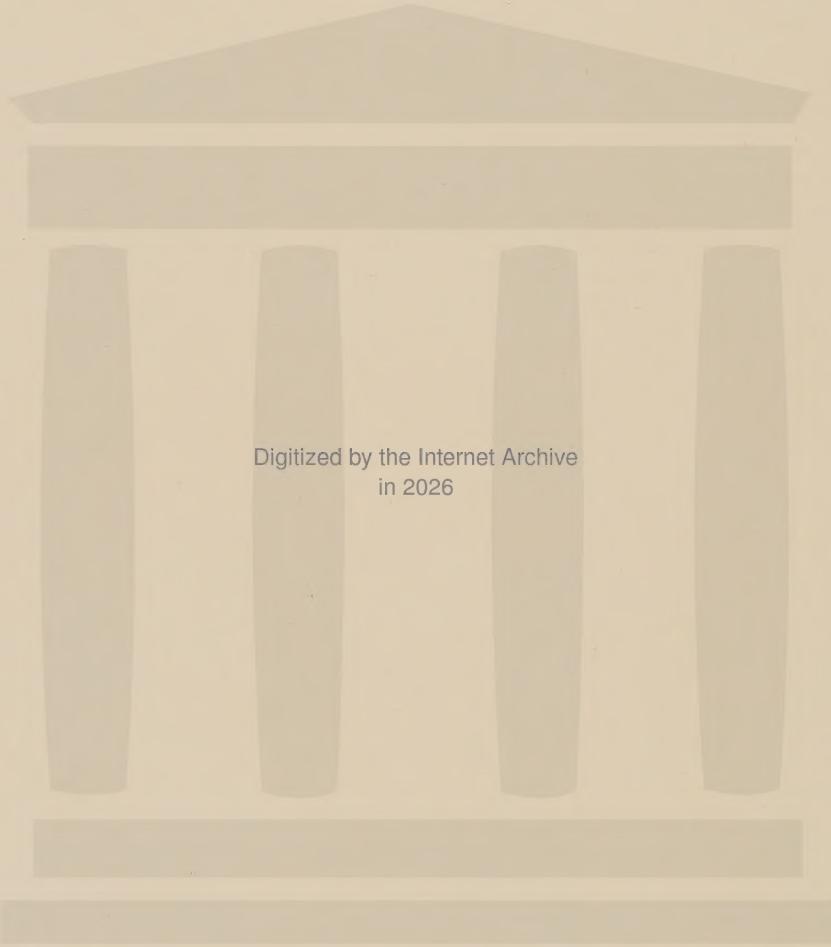
* The method is similar to that used by H. W. Newton and described in *M.N.*, 93, 172, 1933.

† In the case of this eruption, which covered such a small area, visual comparisons with the neighbouring disc must have been peculiarly sensitive to atmospheric seeing; in the u.v. plates this difficulty could be avoided by comparison with other faculae which maintained constant relative brightness *inter se*.



Eruption of 1938 March 31 in continuous light (3220 Å.) and H α . (Scale 1 mm. = 6".7.)
The white arrow points to the nucleus of the eruption, which appears farther from the limb in (4) owing to chromospheric H α .

J. C. Dobbie, W. Moss, and A. D. Thackeray, Two Solar Eruptions.



Digitized by the Internet Archive
in 2026

at $10^h 50^m$ shows a peak brightness more than three times as great as the neighbouring disc, and a total excess intensity of the order of 0.02 per cent. of the normal $H\alpha$ radiation from the whole disc. An attempt has also been made to measure the quantitative increase in the u.v. exposures using the darkening to the limb as a photometric calibration.* The peak intensity was found in this way to be at least 27 per cent. in excess of the neighbouring disc, and the total increase about 0.0014 per cent. of the normal radiation (at 3220 Å.) from the whole disc. A comparison of these figures with the fact that the reflection coefficient of the ionosphere (according to observations by the Field Station) decreased by a factor 10^4 during the eruption indicates that the observed increases in $H\alpha$ and continuous radiation are both secondary effects of some fundamental disturbance to which the ionosphere is peculiarly sensitive. The very close coincidence in time of the two effects suggests an increase in ultra-violet solar radiation rather than a stream of high-speed particles, but it is impossible to say, as yet, whether the source of the disturbance is to be looked for in continuous or monochromatic (e.g. Lyman α) radiation in the ultra-violet. It is excessively unlikely that the continuous radiation from a solar eruption follows a black-body curve; it may arise as emission at the head of the Balmer series. It is hoped to test these two alternatives by observation of future eruptions at different wave-lengths.

Eruption of 1938 April 13 ("The Cricketer").—This eruption was notable for its extended area and length of duration. Ionospheric conditions were disturbed during most of the day, especially during the morning, and the continuous record obtained by the Field Station beginning at $8^h 0^m$ showed practically no trace of an echo on 1.8 megacycles until $12^h 0^m$. A record on long waves (15 km.) showed a number of peaks coinciding with the times of maximum $H\alpha$ activity (see fig. 2). No measure of variation of absorption is available for comparison with our $H\alpha$ intensity curve. No special activity was shown by spectroheliograph observations at $8^h 30^m$ or by spectroheliograms in K and $H\alpha$ obtained at $8^h 26^m$ and $8^h 55^m$ respectively, but at $9^h 14^m$ the area surrounding two large spots near the central meridian was noted as "becoming eruptive." A series of 54 spectroheliograms was obtained by one of us (W. M.) from $9^h 25^m$ to $11^h 12^m$ showing the successive appearances of the area in $H\alpha$ light (Plate 6). A point of great interest which can easily be seen by reference to the plate is that the flocculi retain their general shape (which suggests the title of "The Cricketer") while varying strikingly in their relative intensities during the eruptive period. Attention is drawn in particular to the sudden brightening of three nuclei around the leading spot (to the left in the plate) between $10^h 51^m.9$ and $10^h 54^m.1$ followed by the intensification of an area on the right of this group in the succeeding picture some 90 seconds later. All these areas can be faintly seen on the other pictures, including the first at $9^h 25^m$. No marked radial velocities were noted at the spectroheliograph for any bright area during

* Regions of the solar image unaffected as far as possible by unevenness in the silver coat were carefully selected for this purpose; this source of error does, however, render the photometric result here given only approximate.

the eruption, so that this phenomenon cannot be attributed to the spectroheliograph failing to record some bright regions owing to Doppler shift. Whatever caused the eruption therefore acted as a mechanism which selectively excited existing flocculi.

Fourteen of these spectroheliograms have been measured photometrically by two of the writers (J. C. D. and A. D. T.), and the resulting intensity

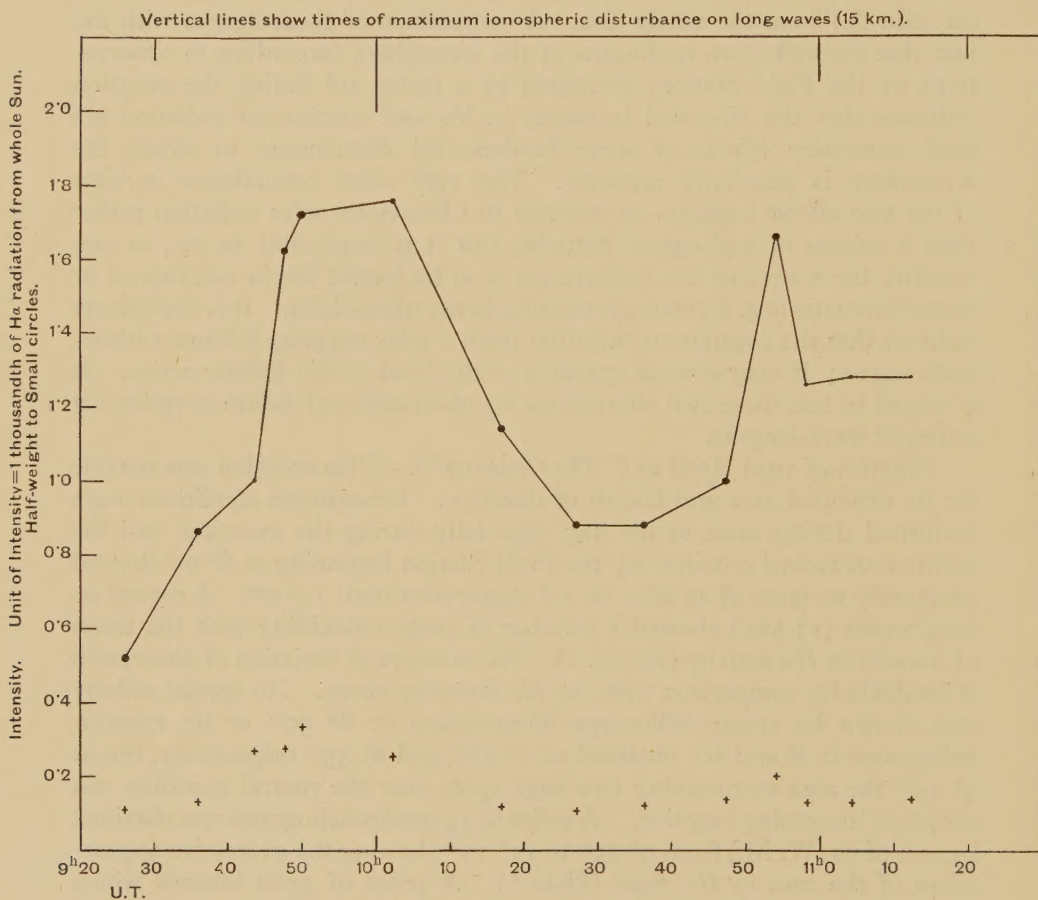
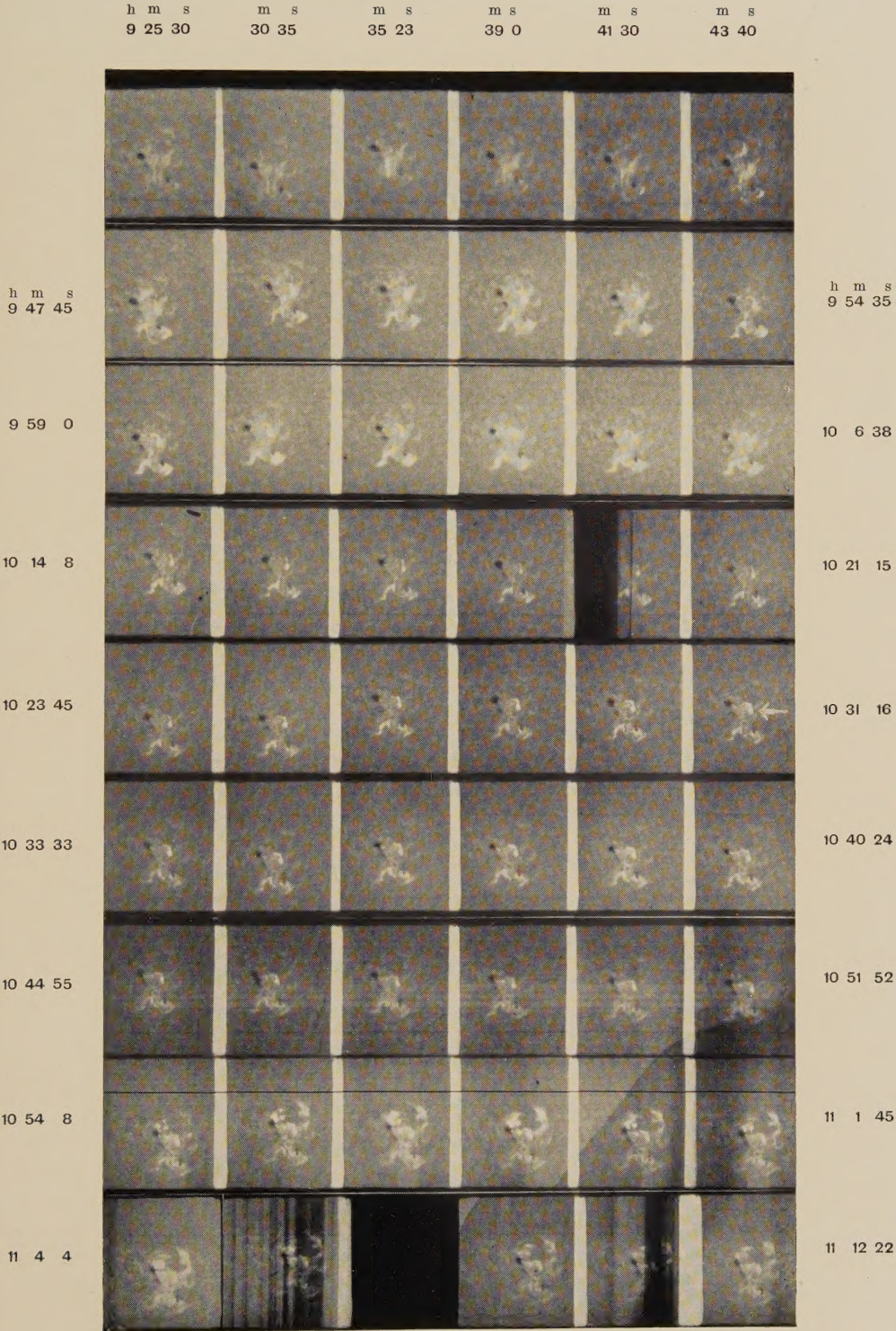


FIG. 2.—Eruption of 1938 April 13.

Crosses refer to contribution from area referred to in the text.

curve is shown in fig. 2 (the ordinate is the total excess intensity over the normal $H\alpha$ radiation from the Sun, the unit being one-thousandth of this radiation). The chief feature of this curve is the double maximum—at about 9^h 54^m and 10^h 54^m—the latter being very short-lived but nearly as intense as the former. Crosses show the contribution from the bright nucleus denoted by an arrow in the spectroheliogram taken at 10^h 31^m 16^s (Plate 6). This region, it will be seen from fig. 2, varied in intensity in roughly the same way as the whole area, contributing at all times about 15 per cent. of the total excess intensity. This behaviour is exceptional, for the other



Solar Eruption of 1938 April 13, in $H\alpha$ light.

(Scale 1 mm.=12".7.)

regions showed only one pronounced maximum, either about $9^h 55^m$ or $10^h 55^m$. It will be noticed that the rise to each maximum is more rapid than the recovery, a phenomenon which is also shown by the eruption of March 31 and by most radio fade-outs. At the primary maximum the total excess intensity amounted to 0.18 per cent. of the total solar $H\alpha$ radiation; this is nine times as great as the figure obtained for the eruption of March 31, a consequence of the more extended area, for the peak intensity of the earlier eruption rose about half a magnitude higher. A quantitative comparison of the ionospheric disturbances associated with these two eruptions is not available, but the long-wave records reveal the highly significant fact that the eruption of March 31, of higher peak intensity, had the greater effect on the ionosphere. Thus in this instance peak intensity of a solar eruption appears to be a more important characteristic than great area. Peak intensities are, however, more difficult to measure than total intensities owing to variable atmospheric definition, as previously noted; it is in fact only possible to give a *minimum* value of peak brightness. The problem is analogous to that of measuring the central intensity of an absorption line with a spectrograph whose resolving power is changing irregularly. The greatest intensity actually measured in "The Cricketer" eruption (in a spectroheliogram taken at $9^h 48^m$ in good seeing conditions) was 3.6 times that of the undisturbed disc.

A series of direct ultra-violet photographs was obtained throughout this eruption, but no marked effects were noticed except for the disappearance of a small spot between $10^h 21^m$ and $11^h 15^m$ in the position of one of the $H\alpha$ nuclei, which brightened at $10^h 53^m$. The fact that the area lay 0.44 solar radii within the limb rendered the surrounding faculæ much less prominent than in the case of the eruption of March 31. In fact the visibility of the eruption of March 31 in continuous light was probably a result chiefly of its very close proximity to the limb.

A comparison of these two eruptions brings out the interesting fact that the first—a short-lived phenomenon with a well-defined maximum—was accompanied by a similarly sharp ionospheric disturbance; the second, a long-lived eruption, extending over a wide area, was accompanied by generally disturbed ionospheric conditions with no strongly pronounced maximum.

Summary.

(1) A description is given of an observational programme designed to obtain (a) photometric intensities of chromospheric eruptions in $H\alpha$ light, (b) evidence of continuous light, especially in the ultra-violet with the aid of a silver filter, accompanying such eruptions.

(2) An eruption, observed 1938 March 31, $12''$ of arc within the east limb, was recorded in both $H\alpha$ and continuous light of wave-length 3220 Å. The observed time of maximum in u.v. light agrees within less than five minutes with that of the ionospheric disturbance recorded by the Field Station of the Cavendish Laboratory on 5 megacycles. The maximum in $H\alpha$ light appears to have been somewhat later. The recovery after maximum

was more rapid in the continuous u.v. than in $H\alpha$. The total increase in $H\alpha$ light amounted to about 0.02 per cent. of the total $H\alpha$ radiation from the Sun; the corresponding increase in continuous light (3220 Å.) was about 0.0014 per cent.

(3) 54 $H\alpha$ spectroheliograms were obtained at approximately 90-second intervals during an extensive eruption observed 1938 April 13. Fourteen of these have been reduced photometrically; the intensity curve shows a pronounced double maximum. The maximum excess intensity in $H\alpha$ amounted to 0.18 per cent. of the total $H\alpha$ radiation from the Sun, *i.e.* nine times greater than the figure for the previous eruption, although the peak intensity was less. The hydrogen flocculi retained their general shape throughout, but varied in their relative intensities; thus the eruption selectively excited existing flocculi. Ionospheric conditions were disturbed throughout, no echo being recorded on 1.8 megacycles by the Field Station during the whole morning. No marked effects were observed in continuous light; this could be attributed to the fact that the area lay much farther within the limb than the previous eruption.

(4) Of the two eruptions, that of March 31, with the smaller area and integrated intensity but greater peak intensity, produced a greater ionospheric disturbance as shown by records on long waves (15 km.).

*Solar Physics Observatory,
Cambridge :
1938 June.*



N.G.C. 4649 and the neighbouring spiral N.G.C. 4647.

(Scale of original plate $1' = 1$ mm., of reproduction $1' = 8$ mm.
4649 has been traced to points outside the boundary of this
picture.)

PHOTOGRAPHIC PHOTOMETRY OF THE ELLIPTICAL NEBULÆ. (SECOND PAPER.)

R. O. Redman and E. G. Shirley.

In continuation of work already described,* photographic observations of nine additional elliptical nebulæ have been made with the 3-foot Common reflector at the Solar Physics Observatory, Cambridge. The results now given are published at this stage since other circumstances prevent further pursuit of the investigation by either author.

Surface Brightness.—The original aim was to obtain photographic magnitudes for the elliptical nebulæ, as far as possible free of systematic or accidental errors greater than $0^m.1$. The method ultimately found most convenient rests on the determination of the surface brightness distribution along the major and minor axes of the nebula as seen in the sky. One of the aims of this paper is to ascertain to what extent observational errors affect this procedure.

The observations and measurements have been carried out almost entirely by one of us (E. S.), following the same method as before (Method 2). During some months of 1936, the tube sensitometer was not available, being removed for use on the eclipse expedition of that year, and for this period plates were calibrated with a wedge, using wave-length $\lambda 4227$. With the tube sensitometer, not only the image of the nebula but also the calibration marks are superposed on the sky fog, which can in consequence be ignored in the course of measurement of the microphotometer tracings. With the wedge this was not possible and the fog had therefore to be measured and allowed for in the reductions. The transmission of the wedge is known from other work† and it should give results quite as accurate as those obtained with the sensitometer.

Ilford Zenith 650 plates have for the most part been used instead of Eastman 40, since the relative superiority of the latter emulsion, noted in our first paper, has not been maintained in more recent batches. Many, probably most, fast plates are liable to show more variation in speed from batch to batch than is commonly supposed, and this, apparently, has been specially true of the particular Eastman 40 samples used by us.

The measured distribution of luminosity on the plate must always be a more or less distorted version of the true distribution. Among the causes of distortion may be enumerated: scattering and irregular refraction of light in the Earth's atmosphere; scattering at the telescope mirror surfaces; finite resolving power of the photographic plate, of the microphotometer and of the telescope; poor guiding. Of these, scattering in the Earth's atmosphere and in the telescope may have an appreciable effect in trans-

* *M.N.*, 96, 588, 1936.

† *M.N.*, 97, 563, 1937.

ferring light over comparatively great distances on the plate. The others are effective mainly over shorter distances and are of secondary importance except near the centre of the nebula, where they entirely dominate the measured light distribution. Guiding errors should be relatively insignificant, as the central regions of the nebula, the most liable to be affected, can be measured on exposures of duration generally not exceeding ten minutes. On long exposures, when guiding errors tend to become greater, and when changes of focus of the telescope primary mirror may also occur, only the outer parts of the nebula are measured, and these are insensitive to this source of trouble.

Neglecting guiding errors, the sum effect of the other factors just mentioned may be assessed by measuring the distribution of luminosity

TABLE I

Average Surface Brightness Distribution in Star Images obtained with the Common 3-foot Reflector

(r = distance from centre of image in minutes of arc, I = surface brightness)

r	$\log I$	r	$\log I$
0.00	2.00	1.2	2.01
.02	1.88	.4	3.91
.04	.61	.6	.82
.06	.31	.8	.74
.08	.03	2.0	.68
.1	0.75	.2	.61
.2	1.72	.4	.54
.3	.23	.6	.48
.4	2.93	.8	.43
.5	.70	3.0	.37
.6	.55	.2	.31
.7	.42	.4	.26
.8	.30	.6	.21
.9	.21	.8	.17
1.0	2.13	4.0	3.13

in star images. If possible, it is best to carry this out on the actual nebula photographs, which in fact was done, a mean distribution being found from two or three plates. However, the procedure was not entirely satisfactory, as the field of good definition on our plates is very small, and, naturally, is occupied by the nebula, so that it was impossible to get a good selection of star images substantially free from coma. Hence special photographs were taken for this purpose, of bright stars on the telescope axis, the exposures never exceeding one minute. On the evening when they were taken there was a certain amount of ground mist present; otherwise the sky transparency was good, but not remarkably so and not better than the average for the nebulae. The telescope mirrors were not in any way specially prepared for these exposures. The star images were measured across north-south and east-west diameters, just as if they had been nebulae, and the final mean radial distribution is given in Table I. In comparison with

the results from star images on nebula plates there is a somewhat sharper peak, and then at intermediate radii, out to $r=1'2$, the intensities are substantially lower, but beyond this become greater and remain so to the farthest point measured ($r=4'$). These differences may be explained by the coma mentioned above and by slightly greater atmospheric scatter in the star plates, the latter not entirely unexpected in view of the state of the sky at the time of observation.

These measures give the observed distribution for a point source. Using them we may compute what would be the observed distribution for any object with a given true distribution. In practice we do not know the true distribution, and the only practicable way of finding it is to work by trial and error from the distortion function and the observed distribution, the complete process being much too long and tedious to be attempted. Our chief concern, however, is merely to make sure that what we observe is mostly determined by the real structure of the nebulae and not by instrumental distortion. Qualitatively this may be verified immediately by direct inspection of the stellar image distribution and comparison with typical nebula images. We see at once that the surface brightness gradient from the centre outwards is much greater in star images than in the nebulae we are measuring, except in the immediate neighbourhood of the centre ($r < 0'1$). It is therefore to be expected that the distortion of the nebula image has been serious only very near the nucleus. Confirmation of a more quantitative nature may be obtained by taking a specimen true luminosity distribution and computing what the observed luminosity would be at various typical points. This has been done, taking for the specimen the observed distribution in N.G.C. 3379. After distortion, $\log I$ at the centre had been lowered by about 0.4; at $r=0'1$ lowered by 0.1; at $r=1'0$ increased by 0.03; and at $r=1'8$ increased by 0.04. This is no proof that the difference between the true and the observed distribution of N.G.C. 3379 was exactly these quantities, but we may expect it to have been of the same order, and considerable support is lent to an earlier judgment, reached solely by experience in measuring the nebula plates, viz. that the effect of observational distortion is great, exceeding say 0.1 in $\log I$, only in regions quite close to the nucleus. Excluding regions for which $r < 0'1$, the only measurement among those published which is likely to be seriously affected in this way is that of the minor axis of N.G.C. 4111, where the gradient of surface luminosity is much greater than usual, the nebula being extremely elongated. Calculations indicate that there is appreciable distortion along the whole of this axis and the figures given must be accepted with some reserve. It is well to note that the integrated brightness of the whole nebula is unaffected.

The distortion just discussed will be present in all astronomical photographs in some degree, however perfect the conditions under which they are taken. While it may be granted that atmospheric conditions can be better than at Cambridge, there remains the finite resolving power imposed by the photographic plate, which at the relative apertures necessary for nebular work is much more important than the true optical resolving power

of the telescope, and this can only be partially evaded by using instruments of longer focal lengths, or other devices for enlarging the scale of the image in the focal plane. It is probable that the determination of the true surface luminosity distribution at the centre of any elliptical nebula is beyond the present resources of astronomy. In this connection, reference may be made to the work of Sinclair Smith on N.G.C. 221.*

The surface brightness results are to be found in Table II, where values are given with the radius as argument, taken along the major or minor axis as the case may be. What was actually measured was the diameter, the nebula being assumed to be symmetrical. This may not be perfectly accurate, but any observed asymmetry, whether real or not, is quite small. Table II also gives the surface brightness range covered in the measurement of each nebula. The centre values, as will be clear from what has been said already, are purely instrumental and in consequence show some correlation with the apparent size and brightness of the whole nebula. (N.G.C. 205, a somewhat abnormal object, stands out as an exception in this regard.) The lower limit of brightness, and also the given diameters, are determined solely by the conditions and method of observation, since, as far as we can tell, these nebulae shade off quite gradually and continuously into empty space.

In judging the accuracy of the measures, we have little but internal checks to guide us, so that, as in most astrophysical work, the errors can only be estimated rather roughly. With regard to accidental errors, the average deviation of observed points from the curve for the major axis of N.G.C. 3610, a case selected at random, is 0.025 in $\log I$ (fig. 1). However, this appears to be better than usual; in the case of the major axis of N.G.C. 3377, where the scatter of the points is somewhat greater than average, the mean deviation is 0.08 (fig. 2). We may suppose that as far as accidental errors are concerned, the greater part of each distribution given in Table II, being taken from a smooth curve through the observations, is correct to a figure substantially better than those just mentioned. The outermost regions are in all nebulae less certain than those nearer the centre, since the observed points are generally fewer, as well as more scattered in $\log I$, so that if the faintest surface brightness is I_r , measures between I_r and $10 I_r$ must be regarded as definitely less reliable than the rest.

Other sources of error have been discussed in the earlier paper and we need only add that further experience has failed to alter the conclusions arrived at there.

Thus as far as we can be guided by internal consistency, the standard of accuracy in these measures appears quite satisfactory. Three nebulae observed by us have also been measured by Hubble.† Of the nebulae in our first paper, only N.G.C. 221 was in Hubble's list, and for it the agreement between the Mt. Wilson and Cambridge measures was fair near the centre, but not very good in the outer parts. Of the nine in the present paper, N.G.C. 3379 and 4649 are among Hubble's nebulae. Comparison of the respective measures in these two cases unfortunately reveals a rather great

* *Ap. J.*, 82, 192, 1935.

† *Ap. J.*, 71, 231, 1930.

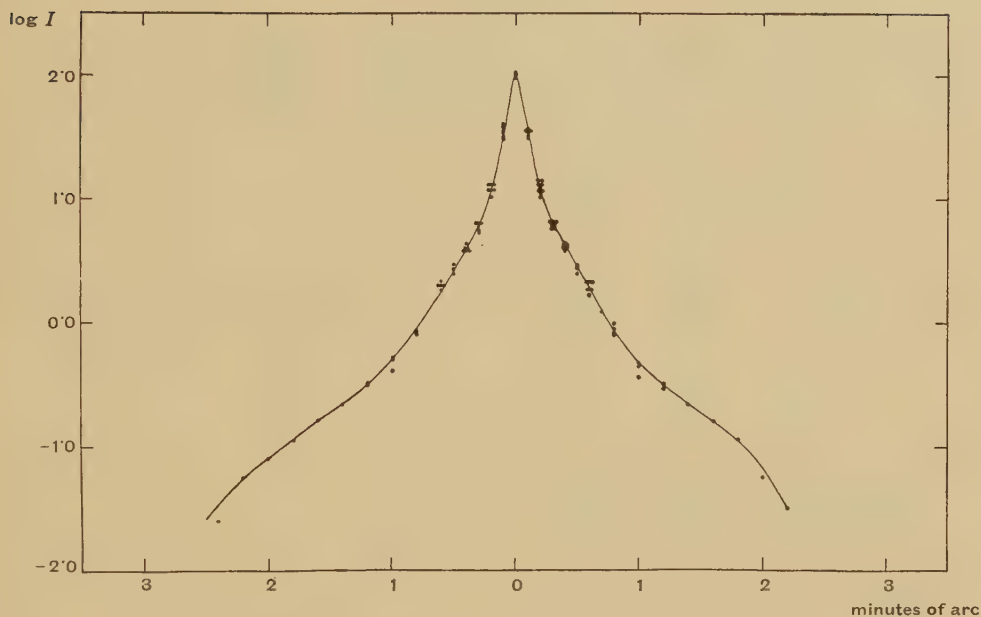


FIG. 1.—Surface brightness distribution along major axis of N.G.C. 3610. The central value, $\log I = 2$, is equivalent to $18^{\text{m.5}}/(\text{second of arc})^2$.

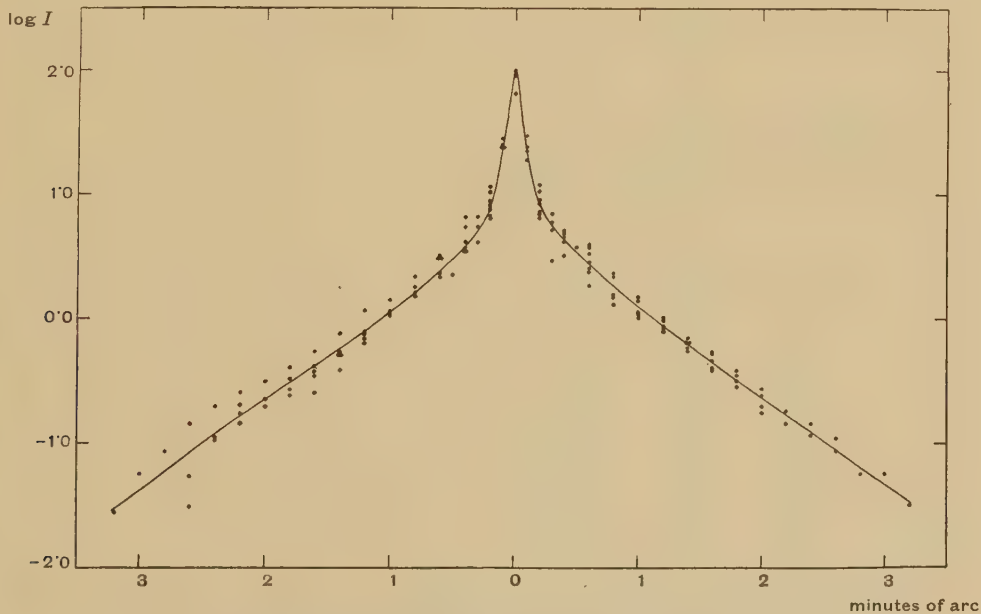


FIG. 2.—Surface brightness distribution along the major axis of N.G.C. 3377. The central value, $\log I = 2$, is equivalent to $17^{\text{m.2}}/(\text{second of arc})^2$.

TABLE II

Distribution of Surface Brightness along Major and Minor Axes of Nebulæ

(Logarithm of surface brightness tabulated against distance from centre in minutes of arc)

1. N.G.C. 205. Type Pec : *

r	Major	Minor	r	Major	Minor
'			'		
0.0	2.00	2.00	1.5	0.79	0.69
.1	1.64	1.62	3.0	.57	.42
.2	.51	.45	5.0	.36	1.80
.4	.35	.25	7.0	.05	.14
.6	.23	.09	9.0	1.70	...
.8	.01	0.94	11.0	.14	...
1.0	0.96	.81			

12 exposures, 3 to 140 minutes.

Surface brightness range, $19^m.0$ to $27^m.1/\text{sec.}^2$.Measures cover diameters $23'.5$ and $14'.0$.

2. N.G.C. 2693. Type E5 :

r	Major	Minor	r	Major	Minor
'			'		
0.0	2.00	2.00	1.0	0.18	1.83
.1	1.53	1.48	.2	.08	.58
.2	.18	.05	.4	1.95	.40
.4	0.78	0.58	.6	.79	.20
.6	.50	.25	.8	.62	...
.8	.32	.04	2.0	.44	...

6 exposures, 6 to 140 minutes.

Surface brightness range, $19^m.2$ to $26^m.3/\text{sec.}^2$.Measures cover diameters $4'.5$ and $3'.4$.

3. N.G.C. 3377. Type E5 :

r	Major	Minor	r	Major	Minor
'			'		
0.0	2.00	2.00	1.4	1.77	1.12
.1	1.34	1.18	.6	.63	2.98
.2	0.89	0.64	.8	.49	.84
.4	.61	.14	2.0	.35	.70
.6	.41	1.84	.4	.07	.42
.8	.23	.61	.8	2.79	...
1.0	.07	.42	3.2	.47	...
.2	1.91	.26			

9 exposures, 5 to 130 minutes.

Surface brightness range, $17^m.2$ to $26^m.0/\text{sec.}^2$.Measures cover diameters $6'.4$ and $5'.3$.* Classifications are according to Hubble, *Ap. J.*, 64, 321, 1926.

TABLE II—*continued.*

4. N.G.C. 3379. Type Eo :

<i>r</i>	Major	Minor	<i>r</i>	Major	Minor
0.0	2.00	2.00	1.2	0.14	0.19
.1	1.55	1.56	.4	.00	.07
.2	.17	.23	.6	1.86	1.93
.4	0.80	0.84	.8	.73	.77
.6	.59	.61	2.0	.57	.60
.8	.43	.45	.2	.39	.41
1.0	.29	.32			

13 exposures, 2 to 230 minutes. Surface brightness range, $17^m.7$ to $24^m.8/\text{sec.}^2$.
 Measures cover diameters $4'.8$ and $4'.8$.

5. N.G.C. 3610. Type E4 :

<i>r</i>	Major	Minor	<i>r</i>	Major	Minor
0.0	2.00	2.00	1.2	1.49	1.46
.1	1.53	1.29	.4	.33	.24
.2	.08	0.80	.6	.19	2.98
.4	0.60	.32	.8	.04	.68
.6	.26	.04	2.0	2.86	...
.8	1.93	1.84	.2	.62	...
1.0	.68	.66			

12 exposures, 5 to 130 minutes. Surface brightness range, $18^m.5$ to $26^m.3/\text{sec.}^2$.
 Measures cover diameters $4'.6$ and $3'.8$.

6. N.G.C. 4111. Type E7 :

<i>r</i>	Major	Minor	<i>r</i>	Major	Minor
0.0	2.00	2.00	1.2	1.84	...
.1	1.55	1.17	.4	.70	...
.2	.15	0.32	.6	.56	...
.4	0.71	1.49	.8	.41	...
.6	.40	2.41	2.0	.25	...
.8	.18	...	.4	2.90	...
1.0	.00	...	.8	.52	...

9 exposures, 2 to 225 minutes. Surface brightness range, $16^m.8$ to $25^m.6/\text{sec.}^2$.
 Measures cover diameters $5'.9$ and $1'.2$.

7. N.G.C. 4649. Type E2 :

<i>r</i>	Major	Minor	<i>r</i>	Major	Minor
0.0	2.00	2.00	1.8	0.33	0.32
.1	1.71	1.71	2.0	.27	.24
.2	.45	.46	.5	.14	.04
.4	.09	.12	3.0	.04	1.86
.6	0.89	0.94	.5	1.94	.68
.8	.75	.81	4.0	.84	.48
1.0	.65	.70	.5	.71	.26
.2	.56	.60	5.0	.57	...
.4	.47	.50	.5	.36	...
.6	.40	.40	6.0	.12	...

8 exposures, 2 to 132 minutes. Surface brightness range, $18^m.3$ to $25^m.4/\text{sec.}^2$.
 Measures cover diameters $12'.0$ and $9'.6$.

TABLE II—*continued.*

8. N.G.C. 5557. Type Eo :

r	Major	Minor	r	Major	Minor
'			'		
0.0	2.00	2.00	1.2	1.42	2.87
.1	1.44	1.27	.4	.24	2.64
.2	0.94	0.81	.6	.08	...
.4	.44	.24	.8	2.93	...
.6	.13	1.82	2.0	.78	...
.8	1.86	.45	.4	.53	...
1.0	.62	.14			

10 exposures, 8 to 170 minutes.

Surface brightness range, $18^m.5$ to $27^m.0$ /sec.².Measures cover diameters $4'.8$ and $2'.9$.

9. N.G.C. 7454. Type E4 :

r	Major	Minor	r	Major	Minor
'			'		
0.0	2.00	2.00	0.8	0.38	1.96
.1	1.68	1.54	1.0	.13	.64
.2	.39	.15	.2	1.88	.33
.4	0.93	0.67	.4	.63	...
.6	.63	.30	.6	.39	...

13 exposures, 10 to 140 minutes.

Surface brightness range $19^m.9$ to $26^m.9$ /sec.².Measures cover diameters $3'.4$ and $2'.5$.

discrepancy, the gradient of luminosity along the axes being much greater in Hubble's figures than our own. We have ascertained that this cannot be ascribed to distortion of the image by scattered light, etc. In one other case, N.G.C. 205, we have measures by Reynolds for comparison.* On the major axis we find rough agreement in the outer parts, but the Cambridge measures show a very much higher intensity at the centre, probably because Reynolds's plates were "burned out" in this region.

Integrated Magnitudes.—To find the brightness of the whole nebula, two steps are necessary: we must first integrate the surface brightness measured and next reduce the sum to the stellar magnitudes scale by comparison with standard stars. We have carried out this integration with the assumption of elliptical isophotal contours. As a test of this method, which can only be approximate, since the contours are not exactly elliptical, the intensity distribution in the image of N.G.C. 3377 was measured on three plates along diameters 0° , 30° , 60° , 90° , 120° and 150° to the major axis, the results enabling one to draw the isophotal contours over a range of diameters $0'.2$ to $4'.4$ (major axis). The total luminosity lying within this area was measured (a) by direct integration, and (b) by assuming elliptical contours and working by the standard method. The results from (b) were

* *M.N.*, 94, 521, 1934.

0^m.06 brighter than from (a). Since 3377 is an E5 nebula, more strongly elongated than the normal object, and this error must decrease with diminishing ellipticity, becoming strictly zero for circular objects, we regard the test as satisfactory, particularly in view of the fact that the aim is to obtain magnitudes with an error not exceeding 0^m.1.

A systematic error may arise from conducting the numerical integration in steps which are too great, since as one proceeds outwards from the centre of the nebula, diminishing surface brightness is always associated with increasing surface area. This has been checked by reintegrating in very small steps, with results never differing from the original integrations by more than 0^m.02.

With regard to the reduction to stellar magnitudes, two extra-focal standard star exposures were usually available. The exposure times for these were kept as short as possible, in order to avoid trouble from a sky of variable transparency. Twenty to thirty minutes were usually sufficient to show a fair amount of the central part of the nebula; enough to measure and relate satisfactorily with the main luminosity distribution, obtained from other photographs made without comparison star fields. In our earlier work, stars in various Harvard Standard Regions were used. For the later measures now being published, magnitudes in the Mt. Wilson Catalogue of Selected Areas have mostly been relied on, and it has been noticed that the internal agreement between star and star is now better than before, although the method of measurement has remained unaltered. The explanation may be that the Harvard magnitudes, based on refractor observations, are affected by colour differences in not quite the same way as the Mt. Wilson magnitudes, which rest on observations with a Newtonian reflector, as does the present work.

Usually three or four stars were used on each plate, but occasionally only two. The average difference in the magnitude of the nebula derived from two independent star plates was 0^m.08, so that the mean of the two should give an average error about 0^m.06. This figure takes no account of errors in the surface brightness distribution, nor, of course, in the magnitudes of the standard stars. Neglecting the last of these two, the total error should remain less than 0^m.1. Where the Harvard Standard Regions have been used, a correction to the international scale of photographic magnitudes is required,* but the necessary data have not yet become available for the fainter comparison stars employed by us. We have used the original uncorrected star magnitudes throughout.†

Table III gives the results from individual star plates, with details as to the particular stars utilized. The surface brightness measures cannot cover the whole nebula, if it has no finite boundary, and so we have made an extrapolation, as explained in the earlier paper, to take account of the outer unmeasured regions. The fraction of the total luminosity depending on this extrapolation is always small and is given in Table III. The final mean magnitudes are in Table IV.

* *H.A.*, 89, No. 1, 1931.

† *H.A.*, 71, No. 4, 1917.

TABLE III

Photographic Magnitudes from Individual Plates

The four columns give respectively the nebula, the comparison stars, the fraction of the integrated luminosity depending on extrapolation of the distribution measures and the resulting magnitude.

N.G.C. 205	S.A. 20, stars 1, 65, 200, 296	0.005	8.74
	" " "		8.63
2693	S.A. 29, stars 73, 2	.056	12.26
	" " 73, 2, 15		12.20
3377	S.A. 79, stars 72, 64, 65	.039	10.83
	" " 72, 64, 65, 32		10.72
3379	S.A. 79, stars 72, 64, 65	.041	10.22
	" " "		10.16
3610	S.A. 13, stars 62, 23	.038	11.51
4111	H.S.R. B5, stars R, U, W, X	.012	11.02
	" " W, X, Y		10.97
4649	S.A. 80, stars 62, 12, 95, 110	.039	9.32
5557	S.A. 59, stars 134, 160, 84, 63	.033	11.95
	" " 84, 63, 14		12.05
7454	S.A. 91, stars 140, 69, 144	.038	12.81
	" " 86, 140, 36, 64		12.88

TABLE IV

Mean Magnitudes

N.G.C.	Cambridge	Shapley-Ames	Holetschek
205	8.69	10.8	8.8
2693	12.23	12.9	12.3
3377	10.78	11.6	10.9
3379	10.19	10.8	9.4
3610	11.51	11.7	11.8
4111	11.00	11.6	10.1
4649	9.32	10.6	9.5
5557	12.00	12.6	12.3
7454	12.85	...	13.3

Discussion.—(i) The present magnitudes, as with those of our first list, tend to be rather brighter than those given by Shapley and Ames in *Harvard Annals*, 88, No. 2, 1932. Excluding the companions to the great Andromeda spiral, N.G.C. 205 and 221, the mean difference Harvard-Cambridge is 0^m.6, while the mean difference without regard to sign is also 0^m.6. It is worthy of comment that the agreement with the Holetschek visual values is on the whole rather better than with the Harvard photographic magnitudes. However, with only a few nebulae available, these results cannot be held to provide any strict check on the statistical errors of the Harvard or Holetschek magnitudes, particularly since these quite possibly depend on the elongation of the nebula image and on the total brightness.

(ii) The diameters to which measures have been taken are considerably higher than in any previous work. The amount of light contained in the outer regions of extremely low surface brightness is surprisingly large, a

fact which must be borne in mind when making photocell observations of these nebulæ. Owing to redistribution of light in the focal plane of the telescope, arising from the causes already discussed, our measures tend to exaggerate the amount of light in the outer parts, but the error at any point seems unlikely to exceed 5 per cent.

(iii) The isophotal contours do not always have a constant ratio of minor to major axis throughout the whole of any one nebula. Instrumental distortion, etc., will always tend to make them circular at the centre, but the effect should be small outside radius $0\cdot1$. The fluctuations of the order of 10 per cent. in the axis ratio which are observed in some nebulæ outside this central zone are probably real, although it is very desirable to have them confirmed by independent observation before any attempt is made to investigate their physical significance. There is no uniform tendency in the changes of shape of the contours, either to become more circular with diminishing surface brightness, or *vice versa*.

There appear to be statistical errors in the axes ratios to be obtained from some catalogues based on eye estimates. For instance, for 15 objects the major/minor axes ratios of Reinmuth's catalogue * are 1.4 times (geometric mean) the Cambridge values, which for the sake of definiteness have been taken at points where the surface brightness is $0\cdot1$ times that observed at the centre. A similar comparison with the Shapley-Ames catalogue (12 nebulæ) gives a factor 1.2. There is no reason to suppose that this is due to error in the present observations. Any "excess" of elongated nebulæ found in statistical counts from such catalogues † must therefore be regarded with suspicion until more careful determinations of the axes ratio are available for large numbers of nebulæ.

Summary.—In continuation of work already published, the distribution of surface brightness has been measured along the principal axes of nine elliptical nebulæ. The amount of instrumental and other distortion has been determined by measuring focal star images, and it is shown that the errors in the surface brightness should be well under $0\cdot1$ in $\log I$, except at distances less than $0\cdot1$ from the centre of the nebula. The surface luminosities have been integrated on the assumption of elliptical isophotal contours, an approximation which has been checked and found sufficiently accurate. Integrated luminosities were reduced to stellar magnitudes by comparison with extra-focal images of standard stars. The results are on the average $0^m\cdot6$ brighter than the values of the Shapley-Ames catalogue. It is noted that the axis ratio of the isophotal contours does not in general remain perfectly constant throughout each nebula, while the visual estimates of Heidelberg and Harvard catalogues appear to give ellipticities systematically too great, a fact of importance in connection with the supposed excess of elongated nebulæ in the sky.

* *Veröff. Sternwarte zu Heidelberg*, 9, 1926.

† Brown, *M.N.*, 98, 218, 1938.

NON-COHERENT FORMATION OF ABSORPTION LINES.

R. v. d. R. Woolley.

1. In Eddington's classical paper, "The Formation of Absorption Lines," * there occurs the passage:

"The crucial question is whether light absorbed in one part of a line is re-emitted in precisely the same part of the line. If so, the blackening in this frequency is independent of what is happening in neighbouring frequencies. . . . For example, if the process is regarded as one of transition between two energy levels, which are not sharp but are composed of narrow bands of energy, the atom is not likely to return to the precise spot in the lower level from which it started and the re-emission will not be the exact reverse of the absorption. In that case the line can only be studied as a whole. Modern attempts to interpret the contours of absorption lines assume (rightly or wrongly) that there is no such redistribution of frequencies."

The last sentence, though written nine years ago, is applicable to-day, because although some writers have shown that they are aware of the difficulty, no progress has been made in developing a theory of the formation of absorption lines taking account of the redistribution of frequency. The difficulties receive prominent treatment in Rosseland's *Theoretical Astrophysics*, where a short account is given of the work of Weisskopf and Wigner on the quantum theory of the mechanism of redistribution. The present writer investigated † the mechanism of redistribution without reference to anything but the principle of detailed balancing, and arrived at expressions which are very similar to Weisskopf and Wigner's. Minnaert wrote a short paper, "Andert sich die Wellenlänge bei der Resonanzzerstreuung?" ‡ in which he pointed out that there was no effect in the case of principal lines. The subject has also been examined by Spitzer.§ We shall make detailed references to his work.

In conformity with Rosseland's terminology, we shall refer to scattering in which there is no redistribution of frequency as coherent scattering, and to scattering with redistribution of frequency as non-coherent scattering.

2. In Eddington's notation the equation of the formation of an absorption line by coherent scattering is

$$\cos \theta \frac{dJ_\nu'(\theta)}{\rho dr} = -(k+l)J_\nu'(\theta) + (1-\epsilon)lJ_\nu' + (k+\epsilon l)J_T.$$

* *M.N.*, 89, 620, 1929.

† *Zs. f. Ap.*, 4, 153, 1932.

‡ *M.N.*, 91, 977, 1931.

§ *M.N.*, 96, 794, 1936.

In the remainder of this paper we shall ignore collisions, and so put $\epsilon = 0$. The above equation will be altered in the non-coherent case only in so far as the re-emission term lJ_{ν}' is concerned. Instead of the re-emission in the frequency ν being equal to the absorption averaged in all directions, it will be equal to the sum of the probabilities that the atoms (in a c.c. of the atmosphere) having absorbed quanta of some frequency ν' will emit one of frequency ν . In fact, if $p(\nu, \nu')d\nu d\nu'$ is the probability that an atom which has absorbed a quantum in the frequency range ν' to $\nu' + d\nu'$ emits one in the frequency range ν to $\nu + d\nu$, the equation of the formation of the absorption line is

$$\cos \theta \frac{dJ_{\nu}'(\theta)}{\rho dr} = -(k + l)J_{\nu}'(\theta) + kJ_T + \int_0^{\infty} p(\nu, \nu')J'(\nu')l(\nu')d\nu'.$$

Following Eddington's standard procedure, this becomes

$$\frac{d^2 J_{\nu}'}{d\tau^2} = 3(1 + \eta_{\nu}) \left[(1 + \eta_{\nu})J_{\nu}' - J_T - \int_0^{\infty} p(\nu, \nu')J'(\nu')\eta(\nu')d\nu' \right].$$

The equations take a slightly simpler form if we write $J(\nu) = J_{\nu}' - J_T$. Then,

remembering that $J_T = H(2 + 3\tau)$ so that $\frac{d^2 J_T}{d\tau^2} = 0$, and also that

$\int p(\nu, \nu')l(\nu')d\nu' = l(\nu)$ from the principle of detailed balancing, the second order differential equations become

$$\frac{d^2 J(\nu)}{d\tau^2} = 3(1 + \eta(\nu)) \left[(1 + \eta(\nu))J(\nu) - \int_0^{\infty} p(\nu, \nu')J(\nu')\eta(\nu')d\nu' \right].$$

This is the fundamental equation for the formation of a single line by non-coherent scattering. Other lines affect the equation through the appearance of fluorescence, into which we shall not go at the present, solely on the principle of dealing with one difficulty at a time. Fluorescence and non-coherence must ultimately be considered together.

It must be noticed in particular that it is assumed that $\eta(\nu)$ is unaffected by any disturbance of the distribution of the atoms in the lower state from the equilibrium value. This stands as a simplifying assumption, but it has a certain correspondence with expectation which makes the case which we consider applicable to some extent to weak lines, namely, that if the lower state is connected with the ground state by a strong line, the radiation in that strong line will recover its equilibrium value in a small depth, compared with that taken by the weak line, so that the disturbance is a skin effect for the weak line.

Spitzer has worked out very carefully the differential equations for the case of an atom having three states, *A* (the ground state), *B* and *C*. The academic case of the single line may be regarded as Spitzer's transition *BC* in the special case where $J(\nu_A^B) = J_T(\nu_A^B)$.

3. We now have to enquire into the form of $p(\nu, \nu')$. According to

Weisskopf and Wigner, the distribution law for the energy of a state whose mean energy is E_k is

$$W_k(E)dE = \frac{A_k}{h} \frac{dE}{\frac{1}{4}A_k^2 + 4\pi^2(E - E_k)^2/h^2},$$

where A_k is the probability of transition from the state, *i.e.* the reciprocal of the lifetime of the state.*

The notation becomes simpler if we write $4\pi\gamma_k$ for hA_k and hx for $(E - E_k)$. Then

$$W_k(x) = \frac{1}{\pi} \frac{\gamma_k}{\gamma_k^2 + x^2}.$$

Now suppose that $J_T A_{jk}$ is the probability that an atom in state j makes an upward transition to state k in thermodynamical equilibrium at a temperature T . Then let $J_T A(x, x') dx dx'$ be the probability that an atom with energy between $E_j + hx$ and $E_j + h(x + dx)$ makes an upward transition to a level with energy between $E_k + hx'$ and $E_k + h(x' + dx')$. Then

$$J_T A(x, x') dx dx' = J_T A_{jk} \frac{1}{\pi^2} \frac{\gamma_k \gamma_j dx dx'}{(\gamma_k^2 + x'^2)(\gamma_j^2 + x^2)}.$$

Now let $E_k - E_j = h\nu_0$ so that ν_0 is the central frequency of the line. Then the probability of the absorption of a quantum in the range ν to $\nu + d\nu$ is given by integrating $J_T A(x, x') dx dx'$ subject to $x - x' = \nu - \nu_0$, so that it is

$$\begin{aligned} J_T d\nu \int_{-\infty}^{\infty} A(x, x + \nu_0 - \nu) dx &= J_T A_{jk} \frac{\gamma_k \gamma_j}{\pi^2} d\nu \int_{-\infty}^{\infty} \frac{dx}{(\gamma_k^2 + (x + \nu_0 - \nu)^2)(\gamma_j^2 + x^2)} \\ &= J_T A_{jk} \frac{d\nu}{\pi} \frac{\gamma_k + \gamma_j}{(\gamma_k + \gamma_j)^2 + (\nu - \nu_0)^2}. \end{aligned}$$

But if $l(\nu)$ is the scattering coefficient

$$\int_{-\infty}^{\infty} A(x, x + \nu_0 - \nu) dx = l(\nu)/h\nu$$

we have accordingly

$$l(\nu) = \frac{h\nu A_{jk}}{\pi} \frac{\delta}{\delta^2 + (\nu - \nu_0)^2},$$

where $\delta = \gamma_k + \gamma_j$ and is the familiar damping constant. For details of the integration carried out above, see Rosseland, *op. cit.*, p. 88. The argument is given by the present writer in *M.N.*, **91**, 981, 1931, and is essentially the same as Spitzer's equation (20).

Now consider the absorption of a quantum of frequency ν' . The probability that this will occur as a result of a transition between the range $E_j + hx$ to $E_j + h(x + dx)$ and the range $E_k + hx + \nu' - \nu_0$ to $E_k + h(x + dx) + \nu' - \nu_0$ is proportional to $W_j(x)W_k(x + \nu' - \nu_0)dx$, and the probability

* Weisskopf and Wigner, *Zs. f. P.*, **63**, 54, 1930; Rosseland, *Theoretical Astrophysics*, p. 87; Spitzer, *loc. cit.*

that this event will occur and will be followed by the emission of a quantum in the range ν to $\nu + d\nu$ is proportional to

$$W_j(x)W_k(x+\nu'-\nu_0)W_j(x+\nu'-\nu)dxd\nu.$$

Integrating this over x gives a quantity proportional to the probability of an absorption in ν' followed by an emission in ν , that is to say to $p(\nu, \nu')l(\nu')$. We have in fact

$$p(\nu, \nu')l(\nu') = h\nu \frac{A_{jk}}{\pi^3 \gamma_j^2 \gamma_k} \int_{-\infty}^{\infty} \frac{dx}{\{\gamma_j^2 + x^2\} \{\gamma_k^2 + (x + \nu' - \nu_0)^2\} \{\gamma_j^2 + (x + \nu' - \nu)^2\}}.$$

(Compare this with *M.N.*, 91, 982, 1931, and Spitzer, *loc. cit.*, equation (19).)

We have further

$$\begin{aligned} \int_0^{\infty} p(\nu, \nu')l(\nu')d\nu' &= h\nu\pi^{-3}A_{jk}\gamma_j^2\gamma_k \int_{-\infty}^{\infty} \int_0^{\infty} \frac{dxd\nu'}{\{\gamma_j^2 + x^2\} \{\gamma_k^2 + (x + \nu' - \nu_0)^2\} \{\gamma_j^2 + (x + \nu' - \nu)^2\}} \\ &= h\nu\pi^{-2}A_{jk}\gamma_j \int_{-\infty}^{\infty} \frac{dx}{\gamma_j^2 + x^2} \cdot \frac{(\gamma_j + \gamma_k)}{(\gamma_j + \gamma_k)^2 + (\nu - \nu_0)^2} \\ &= \frac{h\nu}{\pi} A_{jk} \cdot \frac{\delta}{\delta^2 + (\nu - \nu_0)^2} = l(\nu), \end{aligned}$$

a result which we have already obtained from the principle of detailed balancing.

In the case where $\nu - \nu_0$ is large compared with γ_j or γ_k the integrand in the expression for $p(\nu, \nu')l(\nu')$ takes large values when $\nu' = \nu$ and when $\nu' = \nu_0$, and we have the following approximate results:—

$$\begin{aligned} \int p(\nu, \nu + \xi)l(\nu + \xi)d\xi &= \frac{\gamma_k}{\gamma_k + \gamma_j} l(\nu), \\ \int p(\nu, \nu_0 + \xi)l(\nu_0 + \xi)d\xi &= \frac{\gamma_j}{\gamma_k + \gamma_j} l(\nu), \end{aligned}$$

where it is to be understood that the range of integration is in each case large compared with γ_j or γ_k but small compared with $\nu - \nu_0$. This argument is essentially similar to that used by Spitzer in the derivation of his equation (21).

Comparing with the writer's previous argument in which γ_k was assumed equal to γ_j , if a quantum of frequency ν is absorbed the probability that the energy of the lower state is between $E_j + \hbar x$ and $E_j + \hbar(x + dx)$ is proportional to

$$\frac{dx}{\{\gamma_k^2 + (x + \nu_0 - \nu)^2\} \{\gamma_j^2 + x^2\}},$$

which has maximal values where $x = 0$ and where $x = \nu - \nu_0$. The most probable re-emission is to the central level E_j in either case. If the initial lower state was $x = 0$, the absorption of a quantum of frequency ν is followed by a re-emission of nearly the same frequency; but if the initial energy is

$E_j - (\nu - \nu_0)$, the re-emitted frequency is the central frequency ν_0 . The relative probabilities of these two occurrences are $\frac{\gamma_k}{\gamma_j + \gamma_k}$ (for the re-emission of the same frequency) and $\frac{\gamma_j}{\gamma_j + \gamma_k}$ (for the re-emission of the central frequency). (See fig. 1.)

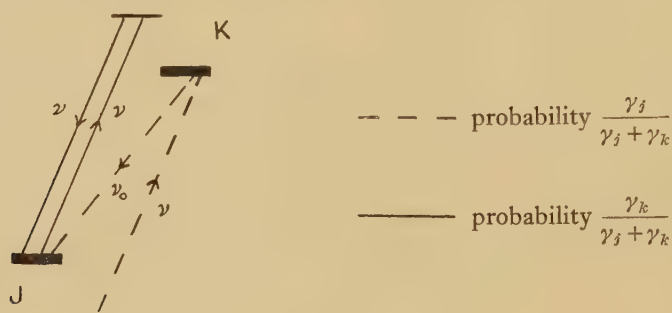


FIG. 1.

4. We have now to solve the infinite set of simultaneous differential equations,

$$D^2 J_\nu = 3(1 + \eta(\nu)) \left[(1 + \eta(\nu)) J(\nu) - \int_0^\infty p(\nu, \nu') \eta(\nu') J(\nu') d\nu' \right],$$

and to make an attack on the problem we resort to a system of successive approximations, the starting-point being the supposition that the radiation in the central frequencies is not much affected by the wings of the line. In order to make the notion of a division between the centre and wings precise, we suppose that there exists a certain frequency range $\Delta\nu$ such that if $|\nu - \nu_0| < \Delta\nu$, $J(\nu) = J(\nu_0)$. We may regard this central range as being created by the Doppler effect (though, of course, the supposition is not an accurate description of the results of the Doppler effect), but clearly the solution of the equations ought not to be contingent on the presence of the Doppler effect. We may regard the non-coherent process itself as tending to create a smoothed region inside which $J(\nu) = J(\nu_0)$.

For definiteness, suppose that we are dealing with a solar line having an intensity 1 on Rowland's scale. Then the value of $\eta(\nu_0)$ is of order 10^3 , and $\bar{\eta}_0$ (the averaged value of the central η as smoothed by the Doppler effect) is about 5. Further, $\Delta\nu = \frac{\nu_0}{c} \sqrt{\frac{2RT}{M}}$ is of order $4 \cdot 10^9$, while δ is of order 10^7 . We shall neglect $\delta/\Delta\nu$ compared with unity, but not $\eta(\nu_0)(\delta/\Delta\nu)$.

Now if $J(\nu) = J(\nu_0)$ for all $|\nu - \nu_0| < \Delta\nu$

$$\begin{aligned} \int_0^\infty p(\nu, \nu') \eta(\nu') J(\nu') d\nu' &= J(\nu_0) \int_0^\infty p(\nu, \nu') \eta(\nu') d\nu' - 2 \int_{\nu_0 + \Delta\nu}^\infty \{J(\nu_0) - J(\nu')\} p(\nu, \nu') \eta(\nu') d\nu' \\ &= \eta(\nu) J(\nu_0) - 2 \int_{\nu_0 + \Delta\nu}^\infty \{J(\nu_0) - J(\nu')\} p(\nu, \nu') \eta(\nu') d\nu'. \end{aligned}$$

For every frequency within the range $|\nu - \nu_0| < \Delta\nu$ we have an equation

$$D^2 J(\nu) = 3(1 + \bar{\eta}_0) \left[(1 + \bar{\eta}_0) J(\nu) - \eta(\nu) J(\nu_0) + 2 \int_{\nu_0 + \Delta\nu}^{\infty} \{J(\nu_0) - J(\nu')\} p(\nu, \nu') \eta(\nu') d\nu' \right].$$

This equation must be averaged over the range $|\nu - \nu_0| < \Delta\nu$. This presents us with a delicate question, whether the average value of $\eta(\nu)$ defined by transfers of radiation is equal to $\bar{\eta}_0$, the average value of the obstruction actually encountered. If we suppose that our flat region $\Delta\nu$ is actually caused by a Doppler effect, it must, strictly speaking, be given by a non-Maxwellian distribution of the velocities, and the obstruction $\eta(\nu)$ is actually constant and equal to $\bar{\eta}_0$ everywhere within the range: on the other hand, we may prefer to picture the actual situation and to regard the η as changing somewhat through the range, so that $J(\nu)$ is not accurately equal to $J(\nu_0)$ everywhere (the $J(\nu_0)$ which we use representing some sort of an average value). In either case the emission η must be averaged for thermal velocities in exactly the same way as the absorption η . We conclude that it is proper to take the average value of $\eta(\nu)$ to be $\bar{\eta}_0$, and we get

$$D^2 J(\nu_0) = 3(1 + \bar{\eta}_0) \left[J(\nu_0) + 2 \int_{\nu_0 + \Delta\nu}^{\infty} \{J(\nu_0) - J(\nu')\} \overline{p(\nu, \nu')} \eta(\nu') d\nu' \right],$$

the bar over $p(\nu, \nu')$ means that an averaged value must be taken.

5. We shall examine only the case where η is constant with τ and take as our zero approximation for $J(\nu_0)$ the coherent value

$$J(\nu_0) = -\frac{2\bar{\eta}_0 H}{1 + \bar{\eta}_0 + \frac{2}{3}q_0} e^{-q_0 \tau},$$

where $q_0^2 = 3(1 + \bar{\eta}_0)$. The equations for the frequencies $|\nu - \nu_0| > \Delta\nu$ are

$$D^2 J(\nu) = 3(1 + \eta(\nu)) \left[(1 + \eta(\nu)) J(\nu) - \int_0^{\infty} p(\nu, \nu') \eta(\nu') J(\nu') d\nu' \right].$$

This is the case where $\nu - \nu_0$ is large compared with γ_j or γ_k , and, as we have already remarked, $p(\nu, \nu') l(\nu')$ is small except where $\nu' = \nu$ and $\nu' = \nu_0$. Inserting the values we have found for $p(\nu, \nu') l(\nu')$ in these two neighbourhoods we find

$$D^2 J(\nu) = 3(1 + \eta(\nu)) \left[(1 + \eta(\nu)) J(\nu) - \frac{\gamma_k}{\gamma_k + \gamma_j} \eta(\nu) J(\nu) - \frac{\gamma_j}{\gamma_k + \gamma_j} \eta(\nu) J(\nu_0) \right]$$

so that if we write β for $\gamma_j/\gamma_k + \gamma_j$ we have

$$D^2 J(\nu) = 3(1 + \eta(\nu))(1 + \beta\eta(\nu)) J(\nu) + \frac{6\beta(1 + \eta(\nu))\eta(\nu)\bar{\eta}_0}{1 + \bar{\eta}_0 + \frac{2}{3}q_0} H e^{-q_0 \tau}.$$

Let $q_\nu^2 = 3(1 + \eta(\nu))(1 + \beta\eta(\nu))$. Then

$$J(\nu) = C(\nu) \cdot e^{-q_\nu \tau} + \frac{6\beta(1 + \eta(\nu))\eta(\nu)\bar{\eta}_0 H e^{-q_0 \tau}}{(1 + \bar{\eta}_0 + \frac{2}{3}q_0)(q_0^2 - q_\nu^2)}.$$

The usual boundary condition $J_v = 2H_v$ at $\tau = 0$ gives

$$C(v) \cdot (1 + \eta(v) + \frac{2}{3}q_v) = -2\eta(v)H \left\{ 1 + 3\beta \frac{\bar{\eta}_0(1 + \eta(v))}{q_0^2 - q_v^2} \cdot \frac{1 + \eta(v) + \frac{2}{3}q_v}{1 + \bar{\eta}_0 + \frac{2}{3}q_0} \right\}.$$

6. The first approximation for $J(v_0)$ is given by

$$D^2 J(v_0) = 3(1 + \bar{\eta}_0) \left[J(v_0) + 2 \int_{v_0 + \Delta v}^{\infty} \{ J(v_0) - J(v') \} \overline{p(v, v')} \eta(v') dv' \right]$$

where now

$$J(v_0) - J(v') = J(v_0) \left\{ 1 + \frac{3\beta \eta(v')(1 + \eta(v'))}{q_0^2 - q_v'^2} \right\} - C_{v'} e^{-q_{v'} \tau}.$$

The remainder of this section is devoted to showing that the terms introduced by the integral may be neglected.

Let $2 \int_{v_0 + \Delta v}^{\infty} \{ J(v_0) - J(v') \} \overline{p(v, v')} \eta(v') dv' = A + B + C$ where

$$A = 2 \int_{v_0 + \Delta v}^{\infty} J(v_0) \overline{p(v, v')} \eta(v') dv',$$

$$B = 2 \int_{v_0 + \Delta v}^{\infty} J(v_0) \frac{3\beta \eta(v')(1 + \eta(v'))}{q_0^2 - q_v'^2} \overline{p(v, v')} \eta(v') dv',$$

$$C = -2 \int_{v_0 + \Delta v}^{\infty} \overline{p(v, v')} \eta(v') C_{v'} e^{-q_{v'} \tau} dv'.$$

To evaluate A , the principle of detailed balancing gives

$$p(v, v') \eta(v') = \overline{p(v', v) \eta(v)}.$$

The bar over $p(v, v')$ means an average over all v inside $|v - v_0| < \Delta v$, so that

$$\overline{p(v, v') \eta(v')} = \frac{1}{2\Delta v} \int_{v_0 - \Delta v}^{v_0 + \Delta v} p(v', v) \eta(v) dv$$

$$= \frac{1}{2\Delta v} \cdot \frac{\gamma_k}{\gamma_k + \gamma_j} \eta(v').$$

We have then

$$A = \frac{J(v_0)}{\Delta v} (1 - \beta) \int_{v_0 + \Delta v}^{\infty} \eta(v') dv'.$$

For $|v - v_0| \gg \delta$ we have $\eta(v) = \frac{\eta(v_0)\delta^2}{(v - v_0)^2}$, and we get

$$A = \eta(v_0) J(v_0) (1 - \beta) \left(\frac{\delta}{\Delta v} \right)^2,$$

which may be neglected in comparison with $J(v_0)$.

By inspection, since $\frac{\eta(v)(1 + \eta(v))}{q_0^2 - q_v^2} < 1$, $B < A$.

C satisfies the inequality

$$|C| < 2 \int_{\nu_0 + \Delta\nu}^{\infty} \frac{1 - \beta}{2\Delta\nu} \cdot \eta(\nu') C_{\nu'} d\nu',$$

or, disregarding the small term in $C_{\nu'}$,

$$\begin{aligned} |C| &< \frac{1 - \beta}{\Delta\nu} \cdot 2H \int_{\nu_0 + \Delta\nu}^{\infty} \frac{\eta^2(\nu)}{1 + \eta(\nu) + \frac{2}{3}q_\nu} d\nu \\ &< \frac{1 - \beta}{\Delta\nu} \cdot 2H \int_{\nu_0 + \Delta\nu}^{\infty} \eta^2(\nu) d\nu \\ &< \frac{2}{3}\eta_0^2 H(1 - \beta) \left(\frac{\delta}{\Delta\nu}\right)^4. \end{aligned}$$

It is not enough, however, to show that $|C|$ is negligible, since it has a different functional form from the other terms in the differential equation.

Regarding $q(\nu)$ as the independent variable, C has the form $\int_A^\infty f(x)e^{-\tau x} dx$,

where $A = q(\nu_0 + \Delta\nu) = q'$, say. But $\int_A^\infty f(x)e^{-\tau x} dx = \phi(A)e^{-\tau A}$ where $\phi(x)$ satisfies $\phi'(x) - \tau\phi(x) = f(x)$. C is accordingly of the form $\mathfrak{B}e^{-q'\tau}$ where $|\mathfrak{B}| < \frac{2}{3}\eta_0^2 H(1 - \beta) \left(\frac{\delta}{\Delta\nu}\right)^4$. The differential solution for $J(\nu_0)$ is now

$$D^2 J(\nu_0) = q_0^2 J(\nu_0) + 3(1 + \bar{\eta}_0)\mathfrak{B}e^{-\tau q'},$$

having the solution

$$J(\nu_0) = \mathfrak{C}e^{-q_0\tau} + \frac{3(1 + \bar{\eta}_0)\mathfrak{B}}{q_0^2 - q'^2} e^{-\tau q'}$$

with the boundary condition

$$\mathfrak{C}(1 + \bar{\eta}_0 + \frac{2}{3}q_0) + \frac{3\mathfrak{B}(1 + \bar{\eta}_0)}{q_0^2 - q_2'^2} (1 + \bar{\eta}_0 + \frac{2}{3}q') = -2\bar{\eta}_0 H.$$

It is now clear that $\mathfrak{B}$ may be neglected, and that the first approximation for $J(\nu_0)$ differs from the zero approximation by terms involving $\eta_0 \left(\frac{\delta}{\Delta\nu}\right)^2$.

7. There is no need to approximate any further, and we can take the results at the end of § 5 for $J(\nu)$. The result for the blackening $r(\nu) = H'_\nu/H$ is

$$r(\nu) = \frac{1 + \frac{2}{3}q_\nu}{1 + \eta(\nu) + \frac{2}{3}q_\nu} - \frac{(1 + \eta(\nu))\eta(\nu)\bar{\eta}_0}{(q_0 + q_\nu)(1 + \eta(\nu) + \frac{2}{3}q_\nu)(1 + \bar{\eta}_0 + \frac{2}{3}q_0)}.$$

The second term is small and *the main difference between the non-coherent and the coherent solutions is the appearance of a quasi-collision term in q_ν in the wings of the line in the non-coherent case*, since we now have

$$q_\nu^2 = 3(1 + \eta(\nu))(1 + \beta\eta(\nu))$$

in the wings. It arises because a fraction β of the light absorbed in any wing frequency is transferred to the central frequency, and it appears that

to a first approximation it does not make much difference whether a certain fraction of the absorbed light is removed by collisions or by non-coherence.

We may repeat here for convenience that $\beta = \frac{\gamma_j}{\gamma_j + \gamma_k}$ where the γ are proportional to the reciprocals of the lives of the states (j lower, k upper).

The importance of collision or quasi-collision terms (such as that used by Stromgren * to represent fluorescence) is very great in connection with the change of contour of the line from centre to limb of the Sun. The existence of such terms in order to explain centre to limb effects has been postulated by H. H. Plaskett † and later by Miss Adam. ‡

Summary.—A solution of the problem of the non-coherent formation of an absorption line, *i.e.* formation of an absorption line in the presence of redistribution of frequency, has been attempted by a method of successive approximation. The main result is the appearance of a strong quasi-collision term in the wings of a weak line. It is suggested that non-coherence is chiefly of importance in centre to limb problems.

* *Zs. f. Ap.*, **10**, 237, 1935.

† *M.N.*, **91**, 870, 1931.

‡ *M.N.*, **98**, 112, 1937.

THE ORIGIN OF SATELLITES.

Raymond A. Lyttleton, M.A., Ph.D.

Introduction.—The theories of the origin of the solar system hitherto put forward have had in common one great difficulty, namely, that in every case the formation of the satellites has not found any definite place in the theory. The mechanism of the formation of planets has undergone a great change, passing from the nebular hypothesis of Laplace to the dynamic encounter theory involving the removal of matter from the surfaces of two stars in a bodily collision with each other, and in each theory the origin of the satellites has been ascribed to obscure and ill-understood causes. The nebular hypothesis fails to give a satisfactory explanation of the existence of the planets so that its satellite difficulties need not be discussed here. The origin of the Moon by tidal resonance is no longer a tenable hypothesis, Jeffreys having shown that the energy would be dissipated as fast as the Sun supplied it, owing to the turbulence in the fluid Earth, when the amplitude of the oscillation of the Earth attained to only a few per cent. of that necessary for disruption. But even if it had been possible to account for the Moon in this way, the origin of all the other satellites would have remained as acute a problem as ever.

When the Sun came to be regarded as the source of the planetary material it was suggested that some of the planets in pursuing highly elliptical orbits might have passed sufficiently near to the Sun for its tidal action to produce satellites. This idea would appear to admit of extension whatever the source of the planets might be, but it can hardly be conceded that the outer planets have ever passed so close to the Sun as tidal action requires. Admittedly, some redistribution of the angular momentum has probably occurred and this might have enabled a particular planet to approach near the Sun, but it is doubtful if the massive Jupiter has ever been as near the Sun as the terrestrial planets are now. Moreover, the satellites of Jupiter do not appear to be the result of this sort of tidal action, which is believed to produce disruption into two or more comparable masses.

Finally, it was proposed by the writer that encounters between pairs of planets, which must at one time have had highly eccentric orbits, might be the explanation of the existence of satellites. But, as Jeffreys has pointed out, encounters of the requisite closeness would take place at prohibitively infrequent intervals, if at all, bearing in mind the present circularity of the orbits, and this consideration alone makes it well-nigh impossible that the satellites, which are so much more numerous than the planets, can be ascribed to such a process.

Evidently there has been no satisfactory theory of the origin of satellites put forward up to now, and the present paper is an attempt to show that a

theory can be developed from a consecutive discussion of the whole situation, beginning with the original collision between two stars, which is supposedly instrumental in generating the planets. It will clarify the argument and lend interest to the subsequent inquiry if mention is made of some of the many remarkable features of which any passably complete theory of the origin of the solar system must give account.

In the first place the masses of the planets show a surprising range, that of Jupiter being several thousand times as great as that of Mercury, as an extreme instance. Again, the ratio of the mass of the Moon to that of the Earth is very much greater than that of any other satellite to its primary—a feature which has often been regarded as indicating that the Moon may have had a wholly different origin from that of other satellites. On the other hand, Jupiter has a mass of order 10^4 times that of any of its four great satellites, which, however, all have masses very similar to that of the Moon.

Previous writers have described the satellite systems as miniature replicas of the planetary system and have inferred that similar processes must have been responsible for their formations. This conclusion may be correct, but there is little ground for making the inference; small bodies moving in the neighbourhood of a massive body must move in orbits around it, and this fact can give no direct information as to the origin of such an arrangement. In point of fact the planetary system differs decidedly from a typical satellite system, in that in the former the angular momentum is segregated almost entirely into the orbital motion of the planets, whereas in the satellite systems practically all the angular momentum is stored in the rotation of the primary. (The Earth-Moon system is excepted of course, though it probably originally conformed to this distinction between planetary and satellite systems.)

Finally, some mention should be made of the mean densities of the planets. As Jeffreys has pointed out in two very important papers,* these densities indicate a difference of composition in the case of the terrestrial planets, which in turn means that in some way the original planetary material had time to separate its heavier and lighter constituents before these planets were formed. Thus the Earth consists of a shell of material, which uncompressed would have a density of about 3.3, surrounding a dense core consisting mainly of iron, bringing the total mean density up to 5.53; whereas the Moon and Mars (and probably Mercury) have mean densities less than 4, indicating that these bodies have little or no heavy core.

These features of the solar system are so pronounced that a theory unable to give some explanation of each of them cannot in any way be regarded as complete.

1. *The Masses of the Primitive Planets.*—One of the chief difficulties of the collision theory of the origin of the solar system has been in explaining how small planets such as Mars and Mercury could have formed in the original filament, which must have had a temperature comparable with that of stellar material. It is possible, as some

* *M.N.*, 94, 823, 1934, and *M.N.*, *Geophys. Suppl.*, 4, 62, 1937.

The present radii and mean densities of the planets are shown in the following table :—

	Radius (10^8 cm. = 1)	Density (Water = 1)
Mercury	0.24	(3.8?)
Venus	0.60	5.21
Earth	0.64	5.53
Mars	0.34	3.95
Jupiter	7.0	1.34
Saturn	5.7	0.70
Uranus	2.5	1.27
Neptune	2.6	1.58

It appears therefore that even Jupiter and Saturn might have some difficulty in forming from material initially at 10^7 degrees. Uranus and Neptune are slightly below the limits found but not alarmingly so ; and it will be seen later how these four planets may have minimised the difficulty. The terrestrial planets and all the satellites are very much below the limit.

Any form of the collision theory is open to this objection if it proposes to produce all the planets in the initial stages. The only way out seemed to be in assuming that the smaller bodies formed from the cooler material of the filament. If, however, the range of temperature is extended to permit the formation of bodies differing as widely in mass as do Jupiter and Mercury, there seems no reason why a slightly greater extension should not be made to include the Moon, and so on. But the Moon cannot now retain an atmosphere, and Mercury shows no trace of atmosphere (though its proximity to the Sun weakens the present argument), so that, as has generally been argued, these bodies can never have been gaseous. Now it seems to the writer that a better plan is to accept the estimate of the initial temperature of the filament between 10^6 and 10^7 degrees and its implications. Some slight range of temperature may be allowed, but there appears to be no escape from the conclusion that the masses of the primitive planets must have been comparable with those of the great planets. This necessarily requires the subsequent formation of the terrestrial planets and the satellites. Accordingly these remarks suggest that the primitive stage of the solar system comprised a number of large massive planets in orbital motion round the Sun. (Any material not gathered into these planets but nevertheless captured by the Sun would of course remain to form the resisting medium.)

2. *The Rotations of the Planets.*—It was Jeffreys who first proposed * that if the origin of the planets was to be attributed to a close encounter between two stars this encounter must have taken the form of a bodily collision. The chief advantage of this process is that the considerable rotations of the planets receive a natural explanation as a result of the shearing motion which would occur as the result of a grazing collision. In addition, it may be mentioned here that the somewhat uncertain action of the tidal forces is dispensed with, or at any rate relegated to a subsidiary part.

* *M.N.*, 89, 636, 1929.

By means of some simplifying assumptions Jeffreys gave a rough calculation of the effects of a collision, and showed that the rotation periods of the primitive planets would be about eight hours if their mean density were the same as the material had when in the stars. Now it is doubtful if the discussion given by Jeffreys would be considered wholly satisfactory, judged by the standards of hydrodynamics of the present day, but there seems no reason to believe that a revised calculation based on modern theory would lead to a result of a different order of magnitude from that given above. In any case, Jeffreys showed in the paper referred to that the primitive planets must be regarded as possessing considerable rotational momentum.

3. *The Development of the Primitive Planets.*—From the point of view of satellite formation the primitive planets must be considered as autonomous bodies, for it seems certain that all outside agencies have been shown incompetent. Now for any planet of the solar system, a simple calculation shows that if its satellites were put into the planet along with their present orbital momenta, the planet would be in no danger of break-up by rotational instability. Even in the extreme case of the Earth-Moon system the stability would not be endangered; this was pointed out long ago by Moulton.

The existence of so-called twins among the planets has often been remarked upon, though the observation has not proved of any heuristic value up to now. Again, Jeans has shown that the fission of a single body would result in a pair of bodies of comparable but necessarily different mass. This investigation by Jeans was conducted mainly in connection with the instability of a rotating star, but it is at least equally applicable to the question of the development of a gradually cooling rotating planet. If these two remarks are taken together it is but a simple step to envisage Jupiter and Saturn, for example, as being the result of the fission of a large primitive planet. The two bodies so formed must, however, be endowed with sufficient energy to escape from each other if planets are ever to have divided in this way, and at this stage a departure would be necessary from those views of the results of fission which have been proposed in connection with the origin of binary stars. Certainly there is nothing impossible about the idea; the element of novelty involved suggests, however, that a careful discussion of the fission problem must be made in order to test the hypothesis of the fission and complete separation of the resulting parts.

4. *The Possibility of Fission of a Planet.*—It has been seen that the primitive planets must have formed from material at high temperature endowed with considerable angular momentum. The process of cooling has been described by Jeffreys,* and it appears that a massive planet would first expand adiabatically to a low density of order 10^{-5} gm. per cm^3 , at which stage liquefaction would begin, the whole process taking a few years.

Now the planet radiates away its energy with negligible loss of angular momentum, and accordingly the mean density $\bar{\rho}$ will increase and so will

* *M.N.*, 89, 731, 1929.

the angular velocity ω , which will be regarded as everywhere the same in the main mass of the planet. The quantity $\omega^2/\bar{\rho}$ also increases, like $\omega^{1/2}$ or $\bar{\rho}^{1/3}$, as is easily seen since the mass and angular momentum remain constant. Stability of a rotating body requires that the quantity $\omega^2/\bar{\rho}$ should not exceed a certain value depending on the density distribution. If the material is concentrated towards the centre a higher value of $\omega^2/\bar{\rho}$ is necessary to produce instability than if the density were uniformly equal to $\bar{\rho}$, but for sufficiently great central condensation instability takes the form of equatorial shedding of matter instead of fission. If G is the constant of gravitation, a body of uniform density ρ loses its stability when its angular velocity is given by

$$\omega^2 = 2\pi G\rho \times 0.142;$$

while if the mass is concentrated almost entirely at the centre instability sets in when

$$\omega^2 = 2\pi G\bar{\rho} \times 0.361,$$

where $\bar{\rho}$ is the mean density of the body.*

It must be assumed that the internal distribution of density was consistent with instability of the fissional type occurring, and this seems reasonable enough for a planet. Accordingly for the purposes of investigation the planet will be regarded as of uniform density, the condition of instability being given by the first of the above equations.

For a uniform gravitating mass, given sufficient angular momentum, stability is completely lost when the configuration reaches the Jacobian ellipsoid whose axes are in the ratios

$$a : b : c : (abc)^{1/3} = 1.886 : 0.815 : 0.651 : 1,$$

where c is the axis of rotation. Once this stage is reached the figure is unstable, and if displaced will elongate in the direction of the longest axis, while a furrow will presumably form in a plane perpendicular to this axis. Thus the fission would appear to take the form of the radial ejection of a part of the body. There seems no reason for regarding this radial velocity as being inappreciably small. Jeans has regarded fission as producing two masses proceeding round each other almost in contact, and in nearly circular orbits, but it does not appear to be essential to follow him in this view. The consideration of this question forms the basis of the succeeding paper on Binary Stars; for the present purposes the fission will be supposed to result in the production of two component masses, and the conditions, as regards energy and angular momentum, for their subsequent escape from each other will be investigated.

5. *The Gravitational Energy of an Ellipsoid.*—The formula for the exhaustion of energy of a body in the form of an ellipsoid in comparison with the state of infinite distension will be required. The result is well known, so that its derivation need not be given. The energy, which is

* Jeans, *Astronomy and Cosmogony* (Camb. Univ. Press, 1929), pp. 219, 254.

of course negative, is with the usual notation

$$-\frac{8}{15}\pi^2 G\rho^2 a^2 b^2 c^2 I,$$

where

$$I = \int_0^\infty \{(a^2 + \lambda)(b^2 + \lambda)(c^2 + \lambda)\}^{-1/2} d\lambda.$$

By means of the substitution $(a^2 + \lambda) \sin^2 \theta = a^2 - c^2$ the integral can be reduced to standard form, thus

$$I = \frac{2}{\sqrt{a^2 - c^2}} F\left(\sqrt{\frac{a^2 - b^2}{a^2 - c^2}}, \cos^{-1} c/a\right)$$

in the notation of elliptic integrals, the modulus being $\sqrt{\frac{a^2 - b^2}{a^2 - c^2}}$. For

the critical Jacobian ellipsoid the value of the integral is found from tables of these functions to be $3.052/a$, so that in terms of the mass and density the gravitational energy of the critical ellipsoid is found to be

$$-0.7825 GM^{5/3} \rho^{1/3}.$$

This may be compared with the value for a sphere, which is

$$-0.9674 GM^{5/3} \rho^{1/3},$$

the numerical factor here being greater since the sphere represents the more exhausted state.

6. *The Energy and Angular Momentum Required for Fission.*—Consider next the energy of rotation which the ellipsoidal form would have at the critical stage. It is

$$\frac{1}{10} M(a^2 + b^2) \omega^2,$$

and since $\omega^2 = 2\pi G\rho \times 0.142$, while the ratios of a , b and c are also known, this energy can be expressed in terms of M and ρ , so that the energy of rotation is

$$0.145 GM^{5/3} \rho^{1/3},$$

and hence the total negative energy of the critical ellipsoid is

$$-0.638 GM^{5/3} \rho^{1/3}.$$

Next suppose that the mass M breaks into two pieces of mass mM and $(1-m)M$ (where $m < 1/2$), both of density ρ , which eventually just escape from each other to infinity. The speeds of rotation of these masses when they are first formed will not differ greatly from ω , but their shapes will not be critical ellipsoids. They can, however, store most energy in this critical configuration, so that an upper limit to their energy would be obtained by supposing them to depart from each other in the critical shape. Since the total energy of two particles in a parabolic orbit is zero, the condition for departure in critical configurations is seen to be

$$m^{5/3} + (1-m)^{5/3} \geq 1,$$

and this cannot be satisfied for non-zero m , since $0 < m < 1/2$. Hence

if the pieces into which a body breaks eventually separate to infinity they will do so as stable bodies (apart from possible further internal contractions).

If it is assumed that the components eventually settle down into bodies of negligible rotation, so that the whole angular momentum is orbital, the corresponding inequality is seen to be

$$0.638 < 0.967\{m^{5/3} + (1-m)^{5/3}\}$$

or

$$m^{5/3} + (1-m)^{5/3} > 0.66,$$

which can be satisfied for mass-ratios less than 1 : 2.

The actual situation must be somewhere intermediate, so that the energy condition may be estimated as

$$m^{5/3} + (1-m)^{5/3} > 0.75, \quad (E)$$

and this can be satisfied for values of $m < 0.25$, corresponding to a mass-ratio of the component of 3 : 1.

The conservation of angular momentum must similarly place a restriction on m , but an additional difficulty is here met with, since the orbital angular momentum is determined solely by the unknown transverse component of the relative velocity of the two bodies, which is assumed to be sufficient for escape. Let it be supposed therefore that the radial and transverse components of velocity are equal at the time when the two bodies first achieve their separate identities, and that they are then moving with escape velocity at a distance apart equal to the sum of their radii when in spherical form. The distance is then underestimated, but the velocity will be overestimated, possibly in two ways: first, the escape velocity is higher; and second, if the masses have not divided at this stage and are going to, presumably there are still some forces at work tending to increase their relative velocity. The orbital angular momentum of the parabolic motion of these two components (supposed spherical) is readily found to be

$$0.79G^{1/2}M^{5/3}\rho^{-1/6}m(1-m)\sqrt{m^{1/3} + (1-m)^{1/3}}.$$

Also the angular momentum of the critical ellipsoid is

$$\frac{1}{5}M(a^2 + b^2)\omega = 0.31G^{1/2}M^{5/3}\rho^{-1/6}.$$

If mM were constrained to spherical form, rotating with angular velocity ω , its angular momentum would be

$$0.15G^{1/2}m^{5/3}M^{5/3}\rho^{-1/6};$$

and this will somewhat underestimate the amount, since the components will not be spherical immediately after fission. A factor 0.20 instead of 0.15 would appear a reasonable value for the coefficient if the pieces are to remain stable. Accordingly the conservation of angular momentum requires

$$0.31 \geq 0.20m^{5/3} + 0.20(1-m)^{5/3} + 0.79m(1-m)\sqrt{m^{1/3} + (1-m)^{1/3}},$$

or

$$1.55 \geq m^{5/3} + (1-m)^{5/3} + 3.95m(1-m)\sqrt{m^{1/3} + (1-m)^{1/3}}, \quad (A)$$

and this can be satisfied for $m \leq 0.21$, corresponding to a mass ratio exceeding 3.6 : 1.

The value of m as determined from either of the inequalities (E) and (A) depends somewhat of course on the estimates made of the factors involving the shapes of the components and the direction of their relative velocity after fission has occurred. It is seen, however, that the conditions cannot be satisfied for equal components. In fact, from the researches of Sir G. H. Darwin * into the configurations of binary stars with uniform density it appears that the angular momentum of the binary star is greater than that of the critical ellipsoid of the same mass for all mass ratios less than 3 : 1. In the present discussion the angular momentum assumed is the same as for a circular orbit, but the energy required is greater if the components are to escape.

The foregoing results indicate that from the point of view of energy and angular momentum requirements there is no reason for dismissing the notion of eventual separation to infinity of the two components resulting from the fission of a rotationally unstable body, but the mass ratio of the components must be regarded in any case as being greater than about 3 : 1.

7. *The Fission of Planets and the Origin of Satellites and the Terrestrial Planets.*—Accepting the view that fission results in two bodies endowed with sufficient velocity for escape, consider the history of a primitive planet from the time of its aggregation onwards.

The initial radius of a planet comparable in mass with Jupiter would be of order 10^7 km. according to Jeffreys,† which is greater than the distances of most satellites from their primaries—another difficulty of the simultaneous formation of satellites and planets. At this stage of high distension the planet would readily be able to store its angular momentum, but the instability would gradually and inexorably be approached as the planet condensed. In this process the liquefaction of materials of high boiling-point, like iron, etc., would occur first and in the cooler outer layers. Droplets of these substances must form and they would then fall towards the centre, so that the materials of high boiling-point would become almost completely separated from the other lighter materials by the very process by which instability is caused. Jeffreys has given conclusive arguments for this process of cooling and separation.‡

When the instability is reached, the mass divides into two main pieces, and it is a fair assumption that a filament of material would find itself streamed out between the two component masses in a similar manner to that in which the planetary material streamed out between the two stars in the original planet-forming encounter. Here, however, there is an important difference in that much of the material has had time to liquefy. The surface material of the planet undergoing fission is at a comparatively

* See Darwin's *Scientific Papers*, III (Camb. Univ. Press, 1910), p. 177, or Jeans's *Problems of Cosmogony and Stellar Dynamics* (Camb. Univ. Press, 1917), p. 134.

† *M.N.*, 89, 731, 1929.

‡ *The Earth* (Camb. Univ. Press, 1929), p. 31.

low temperature, while the inner parts are liquid, so that this stream of matter is in a suitable state to form into a number of small masses, the continued existence of which does not present any difficulty.

In just the same way as the original planetary material escaped from the two parent stars, a portion of this material between the resulting component planets may escape from both these bodies, and it is here that the terrestrial planets may have had their origin. It may be remarked that the combined mass of the terrestrial planets is less than 1 per cent. of that of Jupiter. It is clear that this process gives an opportunity for these small planets to have a difference of composition, and it may be inferred that the layer of cleavage must have been as deep as the outer part of the heavy core of the primitive planet in order that any of this material may have been removed. From this point of view the Earth and Moon must be regarded as coming out of this catastrophe as a pair of separate bodies, though unable to escape from each other, and it therefore appears that it is merely a matter of chance that the Earth possesses a satellite.

It is evident that the end portions of the intermediate material will not escape from the two planets, and some of it may have sufficient transverse motion to enable it to avoid falling back into the adjacent planet. In this way any planet which is the result of fission would come to possess a number of small satellites, and it is clear that they would move near the surface of the planet which captured them, and in the same direction in their orbits as the rotation of their primary. If the initial density of these bodies is assumed to be similar to that of the Moon, the corresponding Roche's limit for Jupiter is at $1.65 \times$ present radius of the planet (this implies a circular orbit; an eccentric orbit would permit closer approach), while for Saturn the smallest distance consistent with stability is 1.48 radii. Accordingly it is possible that some of the captured small bodies could continue in orbital motion around the planets without being disrupted by their primaries. The present densities of the satellites of Jupiter and Saturn appear to be considerably less than that of the Moon, but it is the density of their cores that would be involved in Roche's limit.

This description of the formation of the terrestrial planets and the satellites of Jupiter and Saturn embodies many striking advantages over previous attempts at accounting for them. The difficulty of the formation of small bodies is almost completely removed, while, as far as the writer is aware, these suggestions are the first to be put forward that give any hope of explaining the peculiarities of density distribution of the terrestrial planets and the similarity of the Moon to a number of other satellites in the system.

The axes of rotation of Jupiter and Saturn are in somewhat different directions, but a cataclysmic separation of the kind discussed here might not take place symmetrically. Also, it is not certain that the inclinations of the axes of rotation to the ecliptic have remained unchanged, for the orbital planes are not fixed and may have had considerable inclinations at one time, so that possibly this difference in the directions of the rotation axes is unimportant in the present connection.

8. *The Rotation Period at which Fission Occurs.*—The angular velocity required for instability of a body is proportional to the square root of its mean density, and hence an estimate of the rotation period will not be very sensitive to an inability to state precisely at what density fission would begin. Supposing for simplicity $\rho = 1$, the corresponding rotation period of the critical ellipsoid is about 7.2 hours. (The assumed value of ρ is consistent with the requirement that further instability does not occur.) This period is certainly of the right order of magnitude, but further contraction of the planet would decrease it. The impression one gets is that the rotation periods have been lengthened somewhat to the present values. Tidal friction is not an available agency for this purpose in the case of Jupiter or Saturn, as its effects on the rotations of these primaries cannot be greater than the orbital momenta of their satellites, and this is negligible in comparison with the rotational momentum of the planets. However, if matter has since been added to the planets at their surfaces by some process of accretion, as has been suggested in a different connection by H. N. Russell,* the effect would be in general to lengthen the period of rotation.

A lower limit to the rotation period of the system undergoing disruption can be estimated from considerations of the Earth-Moon system. The Moon was produced in the liquid form and it must have started its career in close proximity to the Earth, but it must always have been outside of Roche's limit. The mean density of the Moon cannot have undergone much change, and the limit for stability depends only on this and the mass of the Earth being given by

$$2.46 (\rho_E/\rho_M)^{1/3} \times \text{radius of Earth.}$$

This may be calculated in terms of the Earth's present radius and density and is about 3 radii. At this distance the Moon would describe a circular orbit in about 7.3 hours. Again, if it were projected at this distance with angular velocity about the Earth of $\sqrt{2}$ times this circular angular velocity, the Moon would escape from the Earth. This would be true *a fortiori* at greater distances, and hence the initial rotation period of the system must have exceeded 7.2 hours. It is seen therefore that any reasonable period of rotation comparable with that of Jupiter indicates that the Moon began its existence as a satellite at a few radii of the Earth distant from it.

9. *The Effect of Central Condensation on Fission.*—It has been seen how considerable separation of the materials may have taken place in the primitive planet from which Jupiter and Saturn were formed before rotational instability was reached. The dividing planet would in such circumstances consist of a denser core surrounded by the lighter materials. Such concentration towards the centre would not necessarily preclude or even hinder fission, though it must occur at a higher angular velocity than if the density were uniformly equal to the average density. The ejected portion would clearly consist mainly of the lighter material

* *The Solar System and its Origin* (Macmillan, 1935), p. 109.

of the original body, for it is this outer material which possesses most of the angular momentum. It is remarkable in this connection that the mean density of Saturn (0.71) is very much less than that of Jupiter (1.34), and it may be surmised that Saturn has little or no heavy core.

10. *Uranus and Neptune*.—It has been shown that Uranus and Neptune are slightly below the limit of radius for a body which could form from material at 10^6 degrees. On the present theory a great planet possessing close satellites must be regarded as being part of an originally greater primitive planet, which goes some way towards resolving the difficulty about the radii; but this difficulty is not acute for these planets. Also, the mass ratio of the two portions into which a planet may divide is less than 1:3, so that Uranus and Neptune cannot have originally been the two components of a single primitive planet. Moreover, their present axes of rotation are in such totally different directions that the bodies from which they were formed cannot have rotated about even approximately parallel axes. There is no difficulty here, however, in explaining what may have become of the companion planets to Uranus and Neptune. For, supposing the primitive planet from which Uranus was formed to have followed its course of evolution more or less homologically to that of the planet from which Jupiter is the larger component, then it is entirely possible that a smaller component ejected by the fission could escape from the solar system. Thus the energy required to eject Saturn from Jupiter to its present mean distance from the Sun could cause Saturn to escape if similarly projected at the distance of Uranus from the Sun. For the energy depends only on the major axes of the orbits and these are at present:

$$\begin{aligned} \text{Jupiter} &= 5.2 \text{ astron. units.} \\ \text{Saturn} &= 9.5 \quad \text{,,} \quad \text{,,} \\ \text{Uranus} &= 19.2 \quad \text{,,} \quad \text{,,} \end{aligned}$$

And clearly

$$\frac{1}{5.2} - \frac{1}{9.5} = \frac{1}{11.5} > \frac{1}{19.2},$$

so that if the fission which produced Jupiter and Saturn occurred at such a distance from the Sun that Jupiter had an orbit of axis 19.2 astron. units, Saturn would have received sufficient energy to escape from the system.

So that the absence of companion planets to Uranus and Neptune is readily explicable in terms of their great distance from the Sun, and furthermore it seems most probable that these remaining planets represent the more massive portions of their respective primitive planets.

Summary.—The masses of the primitive planets must have been comparable with those of the great planets, the terrestrial planets being far too small to have condensed from stellar material. The search for the origin of satellites leads to the consideration of a single planet, extraneous agencies having been shown to be unsuitable for generating satellites. Rotational instability must be involved to produce any disruption of a single mass, but this process produces two bodies of comparable though necessarily

different mass. If Jupiter and Saturn are regarded as the masses resulting from such a catastrophic disruption of a single primitive planet, it must be assumed that the avulsion endowed Saturn with sufficient velocity to escape from Jupiter. A filament of material would stream out between the departing components, and the end portions adjacent to the planet would be captured and possibly give rise to satellites. The intervening portions of the filament would go to neither planet and accordingly themselves become independent planets; the terrestrial planets must be relegated to some such origin as this, if their difference of composition is to be explained, and the existence of the Moon may also find an explanation in this same disruption.

St. John's College, Cambridge :
1938 March 2.

ON THE ORIGIN OF BINARY STARS.

Raymond A. Lyttleton, M.A., Ph.D.

Introduction.—In the foregoing paper it has been seen that the search for the origin of satellites leads to the consideration of the rotational instability of a single body. If the mode of generating satellites that this leads to is to be satisfactory, it is necessary that fission results in the division of the original body into two main masses endowed with sufficient velocity to escape from one another. Now this view of fission is directly contrary to the generally accepted belief which envisages the product of fission as being two masses revolving in close proximity to each other in nearly circular orbits. Accordingly, it is proposed to give here an account of the considerations which have led the writer to dissent from this otherwise unchallenged view of fission.

1. *The Stability of a Rotating System.**—In ordinary discussions of stability the small oscillations of a system are considered without regard to frictional forces. The system is said to be *ordinarily stable* if the principal co-ordinates describing it vary like $e^{i\lambda t}$, while the system is *ordinarily unstable* if some of the co-ordinates vary like $e^{+\lambda t}$, so that the motion does not remain small. But in the case of a rotating system the existence of dissipative action, such as viscosity, within the system itself, may cause it to become unstable in a certain sense before the question of ordinary instability arises. When such dissipative forces are at work the motion of the system may be either of two kinds. First, the amplitude of every vibration of the system may diminish to zero like e^{-kt} , i.e. the co-ordinates are expressible in the form $e^{-kt} \cos pt$, in which case the particular configuration is said to be *secularly* or *permanently stable*. On the other hand, the presence of dissipative forces may cause the amplitude of a vibration to increase like e^{+kt} and the system is then *secularly unstable*; one or more of the co-ordinates varies like $e^{+kt} \cos pt$, and if the system is slightly disturbed from the secularly unstable configuration, it will depart further and further from it. The action of the dissipative forces is continually to increase the amplitude of the oscillations of the system, and if there is no other stable position to which it can settle down, the system must undergo some cataclysm. Accordingly, if a system is ordinarily unstable it is also secularly unstable, but the converse is not necessarily true; a system may become secularly unstable before ordinary instability takes place.

2. *General Discussion of the Fission Problem.*—The configurations of equilibrium of a rotating mass of gravitating liquid of uniform density are associated with the names of Maclaurin and Jacobi. It is well known that

* See Lamb, *Higher Mechanics* (Camb. Univ. Press, 1920), p. 242; Jeans, *Astronomy and Cosmogony* (Camb. Univ. Press, 1929), chap. vii; G. H. Darwin, *Scientific Papers, III* (Camb. Univ. Press, 1910), p. 385.

the series of Maclaurin spheroids remains secularly stable until a certain critical spheroid is reached. Its axes are in the ratios

$$a : b : c : (abc)^{1/3} = 1.197 : 1.197 : 0.698 : 1,$$

the angular velocity and density being related by $\omega^2 = 2\pi G\rho \times 0.187$. At this stage a new series of possible steady configurations comes into existence, the Jacobian ellipsoids. The continuation of the spheroidal series, now more eccentric than the above critical spheroid, still represents ordinarily stable equilibrium forms, but they are secularly unstable. On the other hand, the Jacobian ellipsoids are a secularly stable series of configurations and accordingly are the forms of physical interest. When the ratios of the axes of the ellipsoid reach the values

$$a : b : c : (abc)^{1/3} = 1.886 : 0.815 : 0.651 : 1,$$

the angular velocity and density then being connected by $\omega^2 = 2\pi G\rho \times 0.142$, a new series of configurations become possible forms of equilibrium; these have been generally referred to as the pear-shaped figures. The stability of the Jacobian ellipsoids must be lost at this juncture in accordance with the principle of exchange of stability, and the question arises whether stability passes to the series of pear-shaped figures.

The problem of determining the secular stability of the pear-shaped figure formed the subject of a series of classical papers by Poincare, G. H. Darwin, Liapounoff, and Jeans, to mention only the chief contributors. It was fairly generally expected by Darwin and others that the pear-shaped figure would ultimately prove to be stable, when the analytical difficulties of securing a verdict had been surmounted. Accordingly, when Darwin announced that he had proved the stability, it was believed that the process of generation of binary systems, such as double stars and satellite systems, had been revealed. The initial members of the pear-shaped series differ only slightly from the last stable Jacobian ellipsoid, and this difference may be described as a slight furrowing of the surface in a plane perpendicular to the longest axis, rendering the figure unsymmetrical about the central plane perpendicular to the longest axis.* In other words, a neck begins to form on the body, and if the series were stable the evolution of the mass would consist of a gradual deepening of this neck (as the system slowly changed through physical causes), and it was natural, in view of Darwin's contention as to stability, to envisage the final result as the division of the mass into two unequal components, revolving around each other in practically circular orbits and very nearly in contact.

No sooner had Darwin expressed his final conclusion than Liapounoff announced that an independent investigation, proceeding by a different method from Darwin's, had yielded the result that the pear-shaped figure was unstable. The question was finally settled by Jeans, who by a third method of attack secured a result agreeing with that of Liapounoff and also discovered where Darwin had gone astray, and after eliminating this error showed that the latter's analysis also led conclusively to the result that the pear-shaped figure is secularly unstable.

* See Jeans, *Astronomy and Cosmogony*, p. 219.

The knowledge that the mass becomes secularly unstable at the critical ellipsoidal configuration does not enable the initial motion to be determined unless it is also known whether the mass is ordinarily stable or not. The problem of investigating the ordinary stability has been attacked and solved by Cartan,* who shows that the ordinary stability is lost at the same time as the secularly stability ceases. So that the pear-shaped series is not only secularly unstable but also ordinarily unstable. This conclusion must considerably affect any views about the initial motion of the mass when once the instability is reached.

Darwin's original view that the pear-shaped series represented stable forms, naturally led him to propose the fission theory of the formation of binary stars. When Jeans showed that the series was initially unstable, it was maintained that the departure from the critical ellipsoidal form would not necessarily be rapid, and the fission theory was retained. But this interpretation of the secular instability seems definitely to require that the series be ordinarily stable. This last contribution by Cartan indicates, however, that once the pear-shaped series is reached, the relative motion of the parts is no longer slow and oscillatory, and the motion cannot be traced by a series of steady equilibrium configurations since it must become of a dynamical nature. The instability is of such a kind that the system does not oscillate, but simply departs in an exponential manner from the critical configuration. Though it is not possible to follow out this motion rigorously, it seems entirely possible to advance a number of general considerations which indicate what the result of this instability must be.

As long as the body remains a single mass it cannot get out of this unstable state, for any dissipation of energy in the ensuing motion can only cause the instability to present a more pressing situation than ever to the mass. Accordingly, the body must eventually undergo disruption of some kind, and, on account of the elongated form when the instability sets in, it is natural to suppose that fission into two main masses is the solution which the system adopts to resolve the crisis; this conclusion has never been established, but is generally accepted. It is, however, the subsequent relative motion of the two portions to which especial interest attaches.

If the smaller of the two pieces is denoted by m and the larger by M , then it is clear that fission will result in the part m being ejected away from M with an appreciable radial velocity as well as the transverse component of velocity due to the general rotation of the system. Two cases arise for consideration:

- (1) m may be endowed with insufficient velocity to escape from M .
- (2) m may escape altogether from M .

If (1) occurs, the mass m will proceed to describe an approximately elliptic orbit about M and hence will return to within its original distance from M , but its radial velocity will now be towards M . It is difficult to see how the masses can avoid colliding with one another, and if once a collision occurs there will be further collisions until the masses temporarily re-unite. For

* E. Cartan, *Proc. International Mathematical Congress*, Toronto, 1924, vol. ii (Univ. of Toronto Press, 1928).

the motion of m about M will always pass again through any point of its path, and hence if one collision occurs so must another, and so on, and the consequent dissipation of orbital energy must result in an increase in the depths of these collisions.

It is just possible, however, that m might be ejected in such a way that its radial component of velocity will be exceedingly small when the separate identities of the two pieces first come into being. This assumption would imply that immediately before this stage M and m would be joined by an infinitesimal neck of liquid, and the subsequent absorption of this neck into the two masses can make no sensible difference to their general configuration just after fission, so that m and M might initially come to rotate about one another almost in contact and in almost circular orbits. Thus, if the radial velocity happened to be zero just when division became complete, this result might be possible if the angular velocity were also appropriately adjusted. This is practically the conclusion to which Jeans comes in discussing the process of fission.* But this is not a possible configuration for the following reason. It has been seen that the mass ratio m/M can at most equal about $1/3$ and is probably less. Now Darwin showed that for two masses of uniform density rotating round one another in circular orbits, there is a lower limit to the distance within which the smaller mass may approach the larger in order that it may avoid disruption by the tidal action of the larger mass. (For m/M very small this distance is well known as Roche's limit.) In all cases this distance is about 2.5 radii of a spherical body of mass equal to M and m , and is increased if the mean density of m is less than that of M . Thus, in order to avoid disruption, m would have to be ejected in such a way as to arrive outside of this generalized Roche's limit with at least the circular velocity appropriate to this distance (about 3 radii of M) and have little or no radial velocity. The difficulty about disruption occurs in a more intense form if m and M come near to colliding with one another, so that the possibility of utilizing slight collisions in order to produce circular orbits does not arise, even if it were otherwise a plausible mechanism.

Now these views of the process of fission are not altogether novel. In discussing this question Jeans * comes to very similar conclusions as far as noticing that collisions must take place, but Jeans assumes that the two pieces are of equal mass, and thereby avoids the possibility of disruption, and suggests that the result of the collisions would be to produce a nearly circular orbit with the bodies practically in contact. The assumption of equality of mass is, however, not permissible for there is insufficient angular momentum in the critical ellipsoidal form, and, as has been seen above, the collisions must continue if once one occurs.

Unless these objections can be surmounted there remains only the second possibility (2). If the two components separate with parabolic or slightly hyperbolic velocity, the original mass would have overcome its condition of instability completely and no further difficulties arise: this solution has at least the merit of extreme simplicity. It may, however, be argued that it is improbable that the initial velocity of ejection of m should exceed or equal

* *Problems of Cosmogony and Stellar Dynamics*, p. 252.

the necessary escape velocity ; but if it does not do so, the masses appear unable to avoid colliding and therefore ultimately reuniting, and this would involve further dissipation of energy, and consequently another and more violent attempt at fission, and so on, until the ejected mass does escape completely, the excess angular momentum of the system reappearing as orbital angular momentum.

Accordingly, no objection can be raised against hypothesis (2) and the arguments make clear its claims, though the result cannot be regarded as completely established and is here proposed only as a more satisfactory description of the result of the instability than that represented by the first possibility.

The knowledge that the pear-shaped figure is initially unstable was gained at the pains of great labour, and if only in deference to the prodigious diligence of the investigators, it is essential that the proper interpretation of their conclusions should be made. It seems incredible that the stability is an unimportant factor, which would be so if fission resulted in circular orbits in any case, that is whether the pear-shaped series were stable or not.

3. *The Origin of Binary Stars.*—If the notion of fission resulting in the complete separation of the components is accepted, the present theory of the origin of binary stars must be forfeited. Apart from the doubts raised by the foregoing considerations, it may also be pointed out that the fission theory is unable to explain systems in which the mass-ratio is less than 3 : 1. Similarity of mass appears to be the rule rather than the exception in spectroscopic binary stars, whereas the fission theory would at least predict the converse. Moreover, unless the separations of stars can be regarded as having undergone very great increases, in most cases, the fission theory is capable only of explaining the very closest of binary stars.

The dynamical objections apply equally strongly at whatever stage the stellar material is supposed to undergo fission into two detached masses, and therefore the suggestion that wide binaries may have formed at very low densities, when the stars would have great distension, is open to the present criticisms. Also conclusions relating to the time-scale, as inferred from considerations of the separations and eccentricities of binary stars, can only be made if the initial state of binary star orbits can be predicted theoretically.

Summary.—The pear-shaped series of configurations of a rotating mass of liquid is not only secularly unstable, but also ordinarily unstable, and this indicates that the fission of a body takes the form of the radial ejection of a portion of the mass. Considerations of the subsequent motion of the two components leads to the view that they must be endowed with sufficient energy to escape from each other, otherwise a collision would occur and cause them temporarily to reunite. If this interpretation of fission is correct the origin of binary stars by fission must be forfeited, and any deductions from it relating to the time-scale.

St. John's College, Cambridge :

1938 March 2.

ON THE EVOLUTION OF ECLIPSING BINARIES

Zdeněk Kopal, D.Sc.

Many theoretical discussions on the origin of binaries have been carried out in recent years, but the fission theory has steadily come more into prominence—not because it has succeeded in explaining all the essential data, but rather because the alternative hypotheses, such as the capture theory or that of simultaneous condensation, have been ruled out by the rapidly increasing observational evidence. On the other hand, the fission theory has found a strong support in the study of axial rotation of simple stars. Until recently the observational evidence of a rapid rotation was almost completely lacking, and in the opinions of many astronomers stars had to rotate with equatorial velocities of a few km. per second; but in the light of observations made in the past few years it has appeared that rapidly rotating stars are by no means rare exceptions, and that they are quite common in the early spectral classes A, B and probably O, where the equatorial velocities occasionally exceed 200 or even 250 km./sec.* If we are to accept the fission theory, it is in these types that we should expect the binaries to originate—and a glance at the distribution of eclipsing binaries in the Russell-Hertzsprung diagram† shows that this is indeed the case. Moreover, as was recently pointed out by Kreiken,‡ the components of spectroscopic binaries and eclipsing variables show definitely smaller velocities of axial rotation than simple stars of the same spectral type, suggesting that a part of their initial angular momenta may have been transformed in the orbital momentum of the system—again in agreement with expectations of the fission theory. Thus in the present state of research the fission theory seems to be the only one which is worth considering, and the aim of the present paper will be to examine :

- (i) Whether the fission theory can be reconciled with all essential facts observed in eclipsing binaries.
- (ii) Whether, from this point of view, any of their physical properties suggest an evolutionary progression, or whether they permit a statical interpretation.

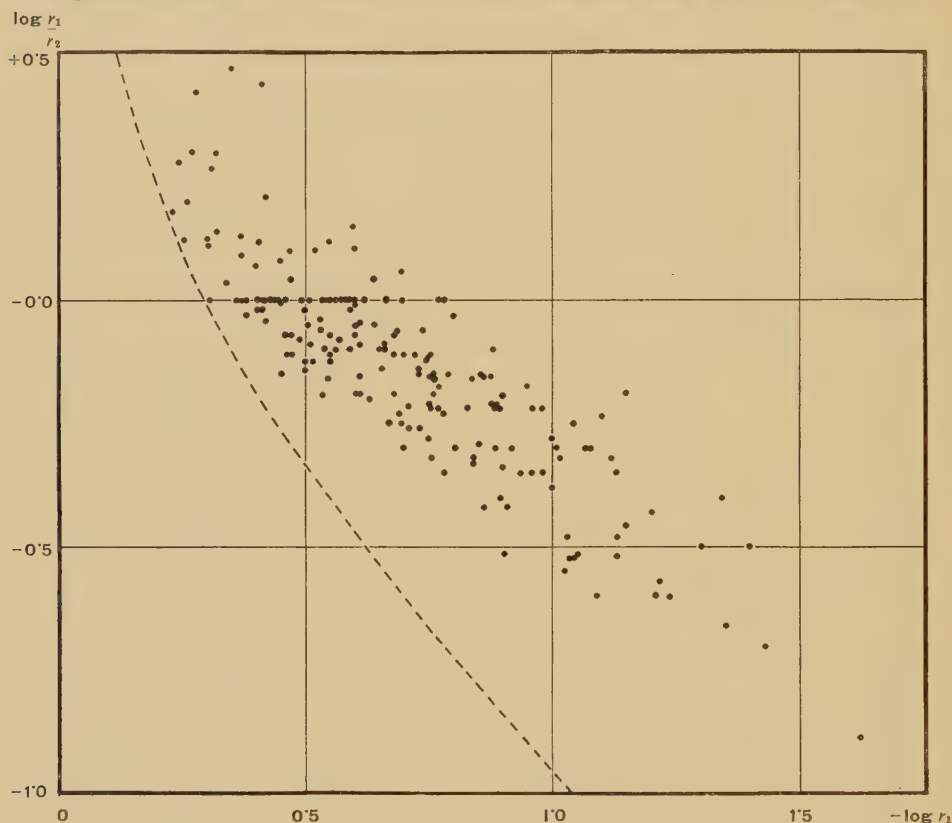
The light curves of eclipsing binaries provide us with the possibility of determining the relative dimensions of the components and the angle of inclination. If the photometric elements can be combined with spectroscopic data, the dimensions, masses and densities can be expressed in absolute units. Up to the present the photometric elements for about

* Cf. O. Struve, *Ap. J.*, **72**, 1, 1930; C. T. Elvey, *Ap. J.*, **71**, 221, 1930; C. Westgate, *Ap. J.*, **78**, 46, 1933.

† Cf. Z. Kopal, *Zeitschrift für Astrophysik*, **9**, 255, 1935.

‡ *Zeitschrift für Astrophysik*, **11**, 323, 1936.

two hundred systems are available, and for about thirty of them the complete sets of absolute dimensions are known. If we know the relative dimensions of a system (expressed in terms of the orbital radius as unity), we can easily, by inverting them, express the orbital radius in terms of the radius of either component, say the primary.* In the diagram this quantity has been



plotted against the radii-ratio $\frac{r_1}{r_2}$, for 198 systems taken (except for a few additions) from Gaposchkin's compilation.† The broken line in the diagram represents the sum of both radii (i.e. $1 + \frac{r_2}{r_1}$), and the distance from it is a measure of the separation of the surfaces.

The diagram shows that as the radii-ratio decreases the separation of the components increases. This applies both to the separation of the centres (indicated by the abscissæ) and the separation of the surfaces (indicated by distance from the broken line). Thus in the closest systems the primary is much larger than the secondary, whilst in the widest systems the reverse is true. If the components were originally in contact the present separation of their surfaces is presumably a rough measure of the

* By "primary" we mean the more massive component in every case.

† *Berl. Bab. Veröff.*, 9, 5, Tab. III, 1932.

stage of evolution reached. We conclude therefore that the radii-ratio decreases as the evolution proceeds. This empirical conclusion agrees with theoretical expectation; because the systems with secondaries larger than primaries could not be produced in their present form by fission.

The exact figures for fissional break-up have been worked out by J. H. Jeans only for two-dimensional series of homogeneous configurations, and here the radii of the components* immediately after the disruption are nearly in the ratio $r_1 : r_2 = 10 : 1$,† and if analogy is any guide, the three-dimensional configurations might be expected to behave similarly. The radii-ratio in which a configuration breaks up will depend on its density condensation; and it is to be expected that with increasing condensation it will rapidly increase, and, as is well known, it tends to infinity for the polytrope $n=5$. The diagram now shows that the actual radii-ratios of

most of the closest binaries are within the limits $1 < \frac{r_1}{r_2} < 10$, which would

point to a nearly homogeneous structure. Since the stars have been considered as appreciably condensed configurations, this has sometimes been used as an argument against the possibility of their fissional break-up, and it has been pointed out that configurations of mean densities as low as those possessed by many eclipsing binaries can scarcely be incompressible. But a number of recent investigations have definitely shown‡ that close eclipsing binaries behave like nearly homogeneous bodies in many different respects, and this seems to turn the argument rather in favour of the fission theory; but it is to be stressed in this connection that low density condensations do not necessarily involve incompressibility.§

Let us turn back to the diagram. It shows that the dimensions of eclipsing binary systems expressed in terms of the radii of their primaries in the interval, say $2 > \frac{r_1}{r_2} > 0.25$, increase about twentyfold. The very important

question arises now to what extent this increase is real, or is due to the changes of our unit, *i.e.* the radius of the primary. The wide separations

of binaries with $\frac{r_1}{r_2} < 0.5$ could *a priori* be considered as due to the contraction

of their components alone, but this would require that the initial densities of the present wide pairs were systematically much lower than those of the present close pairs. Such stars would assume a quite peculiar position in the Russell-Hertzsprung diagram, and the possibility of their existence is therefore very doubtful. On the other hand, the radii and densities of binaries of known absolute dimensions are quite normal compared with those of typical main-sequence stars. This leaves no ground to suspect

* The term "radius" is used only for the sake of simplicity, and denotes the geometric mean of the semi-axes.

† Cf. Jeans, *Problems of Cosmogony and Stellar Dynamics*, pp. 102-116.

‡ Cf. Z. Kopal, *M.N.*, 96, 857, 1936, or the writer's forthcoming article in *The Observatory*.

§ I am indebted to Sir A. S. Eddington for this remark.

that the dimensions of primaries have undergone any more considerable changes * than those due to normal contraction. If so, we are still left with an increase of dimensions of the order 10, and the point of vital importance is to explain this.

So long as a system remains free from external disturbances, its total angular momentum must remain constant. The possibility of external disturbances by close encounters has recently been thoroughly investigated by M. H. H. Walters,† with the result that the effect of possible encounters on the major semi-axes of spectroscopic binaries is practically negligible; and so any increase of orbital dimensions, *i.e.* increase of the orbital momentum, must necessarily be gained at the expense of the rotational momenta of the constituent stars. But, as has been pointed out by Jeans,‡ even in the closest pairs the orbital momentum forms the greatest part of the total angular momentum, leaving very little scope for further increase of dimensions. To obtain a fuller idea about the actual ratio of division of the momentum into its rotational and orbital parts, let us calculate it for a few typical stars.

The total angular momentum M_{tot} is given by

$$M_{\text{tot}} = \sum_{i=1}^2 \mathfrak{M}_i k_i^2 \omega_i + \frac{\mathfrak{M}_1 \mathfrak{M}_2}{\mathfrak{M}_1 + \mathfrak{M}_2} R^2 \omega_1, \quad (1)$$

where $\mathfrak{M}_i$ are the masses, k_i the radii of gyration and ω_i the angular velocities of rotation of the two components, R their distance and ω the angular orbital velocity. As is well known, for a homogeneous body the radius of gyration amounts to two-fifths of its actual radius; inserting this and the Kepler angular velocity for ω and ω_i , we get

$$\frac{M_{\text{orbit}}}{M_{\text{tot}}} = \frac{2}{q + \frac{2}{5}(q+1)(r_1^2 + qr_2^2)}, \quad (2)$$

where r_1, r_2 are the radii of the components expressed in terms of the orbital radius as unity, and

$$q = \frac{\mathfrak{M}_2}{\mathfrak{M}_1}.$$

Inserting appropriate values § in (2) we obtain the following results:—

Star	$\frac{r_1}{r_2}$	M_{tot}	M_{rot}	M_{orbit}
Boss 5070	2.82	1.000	0.109	0.891
VV Orionis	1.84	1.000	0.296	0.704
AG Persei	1.43	1.000	0.078	0.922
WW Aurigæ	1.16	1.000	0.050	0.950
U Ophiuchi	1.00	1.000	0.175	0.825

* The changes of the radii-ratio are probably almost completely due to the abnormal behaviour of the secondaries. Cf. Z. Kopal, *M.N.*, **97**, 646, 1937.

† *M.N.*, **92**, 786, 1932.

‡ *Problems of Cosmogony and Stellar Dynamics*, p. 257.

§ Taken again from Gaposchkin, *loc. cit.*, with the exception of Boss 5070 (G. E. Kron, *Ap. J.*, **92**, 232, 1935) and VV Orionis (H. Schneller, *Kl. Veröff. Berl. Bab.*, **17**, 21, 1936).

As one may easily verify, even under the best conditions, *i.e.* supposing both bodies to be homogeneous, the rotational momentum scarcely amounts to more than one-fifth of the total momentum in very close pairs, and in wider couples the part represented by the rotational momenta is quite insignificant. This implies that no matter how long the tidal friction or other similar tendencies act, the orbital dimensions even of such close pairs as those mentioned above cannot increase more than about twofold—manifestly far too little to account for the feature exhibited in the diagram. If the increase by a factor of the order 10 is real—and we have no way of avoiding it—the ratio $\frac{M_{\text{orbit}}}{M_{\text{tot}}}$ must be much smaller than is given in the

above table. The orbital part of the total momentum depends on directly observable quantities and is therefore fairly well known; the rotational part depends essentially on the density condensation and distribution of the angular velocity in the interior, and is therefore more or less hypothetical. We have already seen that the homogeneous model presents at any rate a good approximation to actual conditions; but as regards the velocity of rotation the situation is less clear. The hypothesis underlying (2) is that both components rotate as rigid bodies with the Keplerian angular velocity. This has been found true for a number of close binaries,* but the observational check (based on the rotational effect) refers only to the surface layers, and the interiors may rotate more rapidly. The increase of the angular velocity inwards has been found inevitable in many recent theoretical investigations;† assuming it, we are free to assign to both components much larger rotational momenta which would be capable of providing for the observed separations. But one point must be stressed. In order that this may be possible, the inward increase of ω must not be accompanied by a large increase of density; or, generally, the density distributions in stellar interiors must be such that *their radii of gyration are not too small*. This, however, strange it may seem, may perhaps be a clue to the actual stellar structures.

Thus, realising that in close binary systems the rotational and orbital momenta may be comparable, there seems to remain no observational fact essentially incompatible with the fission theory. But one interesting feature can still be pointed out. It has already been mentioned at the outset, that if a nearly homogeneous configuration is disrupted by a fissional break-up, the more massive component will always be the larger, the ratio of their volumes being of the order 10. But, as is well known, the mass-ratios are always of the same order—which implies that the secondary will possess higher density than its more massive companion. In extreme cases, such as ζ Aurigæ with the radii-ratio $r_1/r_2=73$, the density ratio ρ_2/ρ_1 is of the order 10^5 .‡ The fission theory in its present state is obviously not able to account for it, because it is confined strictly to homogeneous configurations,

* Cf. P. Swings, *Zeitschrift für Astrophysik*, 12, 70, 1936.

† Cf. Jeans, *Astronomy and Cosmogony*, §§ 249, 257; or Rosseland, *Astroph. Norveg.*, 2, 1936.

‡ Cf. Guthnick-Schneller-Hachenberg, *Abh. d. Preuss. Acad.*, 1935, *Phys.-Math. Kl.*, 1.

but one semi-empirical conjecture can be made. It seems that systems with large values of ρ_2/ρ_1 possess appreciably elliptic orbits; the system already mentioned, ζ Aurigæ, a similar system VV Cephei * and possibly also ϵ Aurigæ,† where ρ_2/ρ_1 are of the order 10^5 – 10^6 , show also the largest eccentricities known among eclipsing binaries. The orbital eccentricities in the closest binaries will no doubt be consequences of the initial velocity of the secondary, and it may be that the order of magnitude of the density ratio is an indication of the centrifugal force with which the secondary was thrown off at the moment of fission. But this conjecture does not yet admit a satisfactory quantitative explanation of observed facts, and, although interesting, is still far from being well established.

If we definitely assume the sense of increasing orbital dimensions as the sequence of evolution, the question may arise whether it is possible to infer from the observed facts anything about the underlying cosmogonic time-scale. Unfortunately no such conclusion can be drawn so long as we have no definite idea how rapidly tidal friction or other similar tendencies (*e.g.* increase of the density condensations of both components) act. It can only be emphasised that there is no observational evidence at present—at least no evidence guaranteed free of selection—which would point to a secular decrease of mass.

The present discussion about the possible course of evolution of close binaries is based mainly on eclipsing variables, because they offer the most complete observational data; but the results will naturally be relevant to all spectroscopic binaries the separations of which do not exceed, say, 30 radii of their primaries. Starting from the pairs that are nearly in contact, there is no trace of any discontinuity in any of their physical properties up to the limit where our observational material ceases. Whether the line of evolution followed so far continues to spectroscopic or even visual binaries, with substantially longer periods and wider separations, is a question on which the present discussion has obviously no bearing.

The eclipsing variables have often been classified as β Lyræ- and Algol-stars, according to whether they display in addition to eclipses light changes due to the elliptical forms of their components. The theory of tidal distortions shows that the mutual forces exerted by the components vary approximately as the inverse third power of the distance. Thus if the increase of their separations is a matter of evolution, the elliptical form of the components, characteristic of the early stages of evolution could not persist as the time elapses, and the β Lyræ-stars would progressively evolve into Algol variables—the two types being successive stages rather than two different groups. Indeed there seems to be no physical reason to support the division of close binaries into groups, and it is probable that they all will follow the same line of evolution. This does not mean that they must follow it individually. The fundamental properties, such as the mass-ratio or initial ratio of radii, will probably depend on the total mass and density condensation of the parent star. The fact that these properties of the large

* S. Gaposchkin, *Harv. Circ.*, 421, 1937.

† Cf. Kuiper-Struve-Strömgren, *Ap. J.*, 86, 570, 1937.

majority of known systems are remarkably uniform suggest that the characteristics of their parent stars were also much alike. The ultimate question why and how the latter acquired their rapid rotation and tendency to disrupt by fissional break-up remains still wholly unanswered.

Summary.—The main subject of the present paper is to investigate from the point of view of the fission theory whether the empirical relation between the radii-ratio and the separation of components in eclipsing binary systems, shown in the diagram and fulfilled remarkably well by nearly two hundred eclipsing systems, can be interpreted as an evolution from close to wide pairs. It is found that the general features agree exactly with the expectations of the fission theory; but in order that the observed increase of dimensions may be possible, the eclipsing systems must have far larger stores of angular momenta in the rotation of their components than the amount calculated from the assumption that they rotate with constant Keplerian angular velocities throughout. A rapid increase of the angular velocity inwards would account for this, and is therefore strongly suggested; but so long as we have no definite idea about the structure of stellar configurations, the observational results do not lend themselves to quantitative conclusions.

In conclusion, it is a great pleasure to express my thanks to Professor Sir A. S. Eddington for stimulating discussions and allowing the very valuable privileges of enjoying the facilities offered by his Observatory, where the above work was carried out. Dr. R. v. d. R. Woolley has again been kind enough to read the manuscript and assist in arranging its final English form.

Cambridge :
1938 June.

Note added in Proof.

In a recent paper * entitled "On the Ellipticity of Close Binaries" W. J. Luyten dealt with certain questions which have bearing on some conclusions of the above paper. But, unfortunately, some of Luyten's inferences cannot be regarded as quite correct. Thus on p. 466 of his mentioned paper Luyten writes: "Since nearly all known cases of apsidal motion are *slower* than would be expected from a reasonable density concentration this would point in the direction of a *decrease* in ω towards the centre." It is a matter of some ambiguity what we call a reasonable density condensation, but if we have in mind, say, the standard model, it should be emphasized that the converse is the case and the observed apsidal motion in nearly all cases is *faster* than that theoretically predicted, suggesting a possible *increase* of ω inwards. Thus Luyten's argument in favour of the capture theory of the origin of binaries turns to the contrary.

On the same page Luyten further states: "If two stars of the same size and similar internal density concentration rotate with the same surface velocity, but if in star A ω increases toward the centre while B rotates as a rigid body, then the observed polar flattening of A will be greater than that of B." It is proper to remark that this inference, although possible, has never been proved rigorously, and every conclusion based on it must consequently be regarded as conjectural.

* *M.N.*, 98, 459, 1938.

A METHOD OF DERIVING THE CONSTANTS OF THE VELOCITY ELLIPSOID FROM THE OBSERVED RADIAL SPEEDS OF THE STARS.

W. M. Smart and S. Chandrasekhar.

(Received 1938 June 3)

1. The usual method of dealing with the radial velocities of the stars in the application of Schwarzschild's ellipsoidal hypothesis involves two preliminary steps: the first is the derivation of the constants of the solar motion from all the observational material, and the second is the removal of the effects of the solar motion from the individual observed radial velocities. The object of these two steps is, of course, to obtain the peculiar radial velocities which form the real observational background of the subsequent analysis. Both the preliminary numerical work and the ensuing calculations involving the peculiar radial velocities are considerable.

In the method to be described we deal with the *observed radial speeds* of the stars, thus avoiding the second step referred to above; there appears also to be a distinct gain in economy of effort as regards the subsequent numerical calculations. Most investigations on the determination of the characteristics of the velocity ellipsoid from radial velocities agree in showing that the lengths of the two shorter axes are not very dissimilar, and in the direct method of applying Schwarzschild's hypothesis to proper motions equality is generally assumed; accordingly, in the sequel we make the assumption that the two shorter axes are actually equal.

2. We assume that the velocity ellipsoid is given by

$$K^2U^2 + H^2(V^2 + W^2) = 1, \quad (1)$$

where (U, V, W) are the components of the peculiar velocity of a star referred to the principal axes of the ellipsoid and K is less than H ; thus the U -axis is the axis of greater mobility.

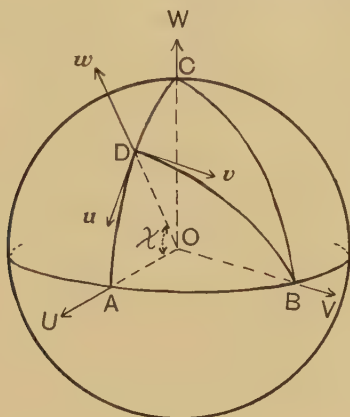
Consider a small region of the sky at D , OD making an angle χ with OA (the U -axis). Owing to the form of (1)—the section of the spheroid by the plane $U=0$ is a circle—we can choose the W -axis to lie in the plane of OA and OD ; the V -axis is then specified as in the figure.

Let N be the number of stars in the region at D ; then the number with components of peculiar velocity between (U, V, W) and $(U+dU, V+dV, W+dW)$ is

$$\frac{NKH^2}{\pi^{3/2}} e^{-K^2U^2 - H^2(V^2 + W^2)} dUdVdW. \quad (2)$$

Let (u, v) be the rectangular components of the tangential peculiar

velocity at D and w the peculiar radial velocity; we take u to be in the plane AOD and v in the plane BOD . The relations between the components



(U, V, W) and the components (u, v, w) are easily seen from the figure to be

$$\left. \begin{aligned} U &= u \sin \chi + w \cos \chi \\ V &= v \\ W &= -u \cos \chi + w \sin \chi \end{aligned} \right\} \quad (3)$$

Inserting these expressions in $K^2 U^2 + H^2 (V^2 + W^2)$ we find, denoting the resulting function by $f(u, v, w)$, that

$$f(u, v, w) = u^2(K^2 \sin^2 \chi + H^2 \cos^2 \chi) + H^2 v^2 + w^2(K^2 \cos^2 \chi + H^2 \sin^2 \chi) - 2uw \sin \chi \cos \chi (H^2 - K^2). \quad (4)$$

Also,

$$dUdVdW = \frac{\partial(U, V, W)}{\partial(u, v, w)} du dv dw = du dv dw.$$

Hence by (2), the number, dN , of stars with components of peculiar velocity between (u, v, w) and $(u + du, v + dv, w + dw)$ is given by

$$dN = \frac{NKH^2}{\pi^{3/2}} e^{-f(u, v, w)} du dv dw. \quad (5)$$

We write (4) in the form

$$f(u, v, w) = (Ju - Aw)^2 + H^2 v^2 + L^2 w^2, \quad (6)$$

in which

$$J^2 = K^2 \sin^2 \chi + H^2 \cos^2 \chi,$$

and it is easily seen that

$$L^2 = \frac{H^2 K^2}{J^2};$$

we do not require to specify the coefficient A .

It follows that the number, dn , of stars with peculiar radial velocities between w and $w + dw$ is given from (5) and (6) by

$$\begin{aligned} dn &= \frac{NKH^2}{\pi^{3/2}} e^{-L^2 w^2} dw \int_{-\infty}^{\infty} e^{-(Ju-Aw)^2} du \int_{-\infty}^{\infty} e^{-H^2 v^2} dv \\ &= \frac{NHK}{J\pi^{1/2}} e^{-\frac{H^2 K^2}{J^2} w^2} dw, \end{aligned}$$

or, writing

$$j = \frac{HK}{J} = \frac{HK}{(K^2 \sin^2 \chi + H^2 \cos^2 \chi)^{1/2}}, \quad (7)$$

we have

$$dn = \frac{Nj}{\pi^{1/2}} e^{-j^2 w^2} dw. \quad (8)$$

The mean peculiar radial speed, $\bar{w}$, is given by

$$\frac{1}{2} N \bar{w} = \int_0^{\infty} w dn,$$

from which

$$\bar{w} = \frac{1}{j\pi^{1/2}}.$$

This result shows that $1/j$ has the dimensions of speed.

3. We assume now that the constants of the solar motion are known. Let λ be the angular distance of the region at D from the solar antapex and let S be the solar speed. The observed radial velocity, ρ_0 , will then consist of the peculiar radial component, w , and the parallactic component, ρ , given by

$$\rho = S \cos \lambda, \quad (9)$$

so that

$$\rho_0 = w + \rho.$$

We denote by R_0 the observed radial speed of a star, so that

$$R_0 = |\rho_0|,$$

and we shall find an expression for the mean radial speed of the N stars in the region at D by an analysis similar to that previously given* by one of us.

Of the N stars let there be $2n_1$ stars for which $0 \leq |w| \leq \rho$ and $2n_2$ stars for which $|w| > \rho$. (We assume that ρ is positive; the argument is similar if ρ is negative.)

Then, (i), there are $n_1 + n_2$ stars for which ρ_0 is positive for all positive values of w , so that

$$|\rho_0| = |w| + \rho,$$

from which

$$\sum_{n_1+n_2} R_0 = \sum_0^{\infty} |w| + (n_1 + n_2)\rho. \quad (10)$$

* W. M. Smart, *M.N.*, 95, 121, 1934.

Also, (ii), there are n_1 stars for which ρ_0 is positive, with w negative but numerically less than ρ . Hence

$$|\rho_0| = -|w| + \rho,$$

so that

$$\sum_{n_1} R_0 = - \sum_0^{\rho} |w| + n_1 \rho. \quad (11)$$

Again, (iii), there are n_2 stars for which ρ_0 is negative, with w less than $-\rho$. Accordingly,

$$-|\rho_0| = -|w| + \rho,$$

so that

$$\sum_{n_2} R_0 = \sum_{\rho}^{\infty} |w| - n_2 \rho. \quad (12)$$

Adding the results in (10), (11) and (12) we obtain

$$\sum_N R_0 = 2 \sum_{\rho}^{\infty} |w| + 2n_1 \rho. \quad (13)$$

Now, from (8),

$$n_1 = \frac{Nj}{\pi^{1/2}} \int_0^{\rho} e^{-j^2 w^2} dw,$$

and

$$\begin{aligned} \sum_{\rho}^{\infty} |w| &= \frac{Nj}{\pi^{1/2}} \int_{\rho}^{\infty} w e^{-j^2 w^2} dw \\ &= \frac{N}{2j\pi^{1/2}} e^{-j^2 \rho^2}. \end{aligned}$$

Also, denoting the mean value of R_0 for the N stars by $\bar{R}_0$, we have

$$\sum_N R_0 = N \bar{R}_0.$$

The formula (13) then becomes

$$\begin{aligned} \bar{R}_0 &= \frac{1}{\pi^{1/2}} \left\{ \frac{1}{j} e^{-j^2 \rho^2} + 2j\rho \int_0^{\rho} e^{-j^2 w^2} dw \right\} \\ &= \frac{1}{j\pi^{1/2}} \left\{ e^{-j^2 \rho^2} + 2j\rho \int_0^{j\rho} e^{-x^2} dx \right\}, \end{aligned}$$

which we write in the form

$$\bar{R}_0 = \frac{1}{j\pi^{1/2}} F(\xi), \quad (14)$$

where

$$F(\xi) = e^{-\xi^2} + 2\xi \int_0^{\xi} e^{-x^2} dx, \quad (15)$$

and

$$\xi = j\rho = jS \cos \lambda. \quad (16)$$

Since, by section 2, $1/j$ has the dimensions of speed, ξ is a pure number.

The function $F(\xi)$ has been previously tabulated.* It is clearly an even function of ξ , so that $F(-\xi) = F(\xi)$.

Using (16) we write (14) in the form

$$\frac{\pi^{1/2} \bar{R}_0}{S \cos \lambda} \xi = F(\xi). \quad (17)$$

This is our fundamental formula.

It is to be remarked that if, in the preceding analysis, ρ is negative—that is to say, $90^\circ < \lambda < 180^\circ$ —we replace $\cos \lambda$ in (17) by its numerical value.

4. For a given region we suppose the value of R_0 to be found from the observed radial velocities, and, since S and λ are presumed to be known, the coefficient of ξ on the left of (17) is accordingly obtained. We can then solve (17) for ξ by a simple graphical process, or by any other suitable method. Thus we obtain the appropriate value of ξ for each region of the sky. Now, by (7) and (16),

$$\xi = S \cos \lambda \frac{HK}{(K^2 \sin^2 \chi + H^2 \cos^2 \chi)^{1/2}},$$

which we write in the form

$$\left(\frac{S \cos \lambda}{\xi} \right)^2 = \frac{1}{H^2} + \left(\frac{1}{K^2} - \frac{1}{H^2} \right) \cos^2 \chi. \quad (18)$$

It is clear from this formula that $1/K$ and $1/H$ have the dimensions of speed.

The left-hand side of (18) is now known for a given region; denote it by C , and let

$$\frac{1}{H^2} = p, \quad \frac{1}{K^2} - \frac{1}{H^2} = q. \quad (19)$$

Then (18) becomes

$$p + q \cos^2 \chi = C. \quad (20)$$

We assume,† in accordance with the results of many investigations, that the axis of greatest mobility (the U -axis) lies in the galactic plane; let its galactic longitude be G_0 . If (G, g) denote the galactic co-ordinates of the region, then

$$\cos \chi = \cos g \cos (G - G_0), \quad (21)$$

so that (20) becomes

$$p + q \cos^2 g \cos^2 (G - G_0) = C. \quad (22)$$

This is the equation of condition from which we have to determine the values of the three unknowns p , q and G_0 . When these have been found, the semi-axes $1/K$ and $1/H$ of the velocity spheroid are then obtained from (19); in particular, the ratio of the axes of the velocity spheroid, namely K/H , is $\left(\frac{p}{p+q} \right)^{1/2}$.

* W. M. Smart, *M.N.*, 96, 167, 1936.

† The assumption is not necessary for a complete analysis on the lines of the present investigation.

5. For the determination of p , q and G_0 by a least-squares solution we distinguish two cases ; in the first, all the stars under investigation lie within a narrow zone of galactic latitude, and in the second case they may be distributed over the sphere in any way.

Case 1.—In (22) put

$$q_1 = q \cos^2 g.$$

Then the equation can be written

$$P + Q \cos 2G + R \sin 2G = C, \quad (23)$$

where

$$\left. \begin{aligned} P &= p + \frac{1}{2}q_1 \\ Q &= \frac{1}{2}q_1 \cos 2G_0 \\ R &= \frac{1}{2}q_1 \sin 2G_0 \end{aligned} \right\}. \quad (24)$$

(23) is now the equation of condition, and P , Q and R can be found from the resulting normal equations ; p , q_1 and G_0 are next found by means of (24) and then q is obtained ; the values of the semi-axes, $1/K$ and $1/H$, of the velocity spheroid are finally determined by means of (19).

Case 2.—The equation (22) cannot now be put directly into a form suitable for the application of the least-squares method. We proceed as follows : Let p_0 , q_0 and G_0' be approximate values of p , q and G_0 —we may, for example, suppose these approximate values to be the corresponding values obtained, by the method indicated in Case 1, for stars lying within a particular zone of galactic latitude, or we may take as the approximate values the results derivable from other investigations.

Let C_0 be the value of

$$p_0 + q_0 \cos^2 g \cos^2 (G - G_0')$$

for a given region. Then, writing $p = p_0 + \Delta p$, $q = q_0 + \Delta q$ and $G_0 = G_0' + \Delta G_0$, we find easily that

$$\Delta p + \frac{C_0 - p_0}{q_0} \Delta q + 2(C_0 - p_0) \tan (G - G_0') \Delta G_0 = C - C_0. \quad (25)$$

This is now the equation of condition ; Δp , Δq and ΔG_0 are then obtained from the resulting normal equations, and, finally, we obtain G_0 and the semi-axes $1/K$ and $1/H$.

Summary

A method is developed, based on Schwarzschild's hypothesis, for the determination of the constants of the velocity ellipsoid (with two equal axes) from the observed radial *speeds* of the stars.

1938 May 23.

University Observatory,

Glasgow, W. 2.

Yerkes Observatory, Williams Bay,

Wis., U.S.A.

THE IONOSPHERIC ECLIPSE OF 1940 OCTOBER 1.

H.M. Nautical Almanac Office.

(Communicated by the Astronomer Royal)

The path of totality of the total solar eclipse of 1940 October 1 has already been given in *M.N.*, 93, 181-184, and will be published in greater detail, and with more definitive corrections to the positions of the Sun and Moon, both in the national ephemerides and in a separate Eclipse Supplement to be issued shortly by the U.S. Naval Observatory. This paper gives the projections on the Earth of the paths of totality of the optical eclipse at heights of 100 kilometres and 300 kilometres in the ionosphere, corresponding to the ionized regions E and F in the upper atmosphere, from which wireless waves are reflected.

There is still considerable divergence of opinion as to the agency responsible for the ionization of these regions, and a total solar eclipse provides the most valuable opportunity for the differentiation of the possible causes. In particular, the abnormal Region E reflections may be caused in a number of ways, and it has been suggested * that the responsible agency might possibly be determined by using observations at the time of solar eclipse to find whether it travels with the speed of light or not. The normal ionization of Region E is due to ultra-violet light, but the evidence from which the reason for the abnormality can be found is still inconclusive. Further, opinions differ as to the ionizing agency for Region F, in particular for Region F₂, and here again valuable information can be obtained by the observations of signal strength during an eclipse. It is hoped that the tracks of totality given here will provide, for the optical eclipse, the material data upon which such observations can be based. It has been suggested † that the speed of the corpuscles that may be the cause of the ionization is of the order of 1000 miles a second, and in this case the track will be quite different from that of the optical eclipse, and will occur about two hours earlier.

The heights considered of 100 and 300 kilometres are too large for the application of the relatively simple differential formulæ from the path of totality on the Earth. For the central line of the eclipse, however, it is possible to base the calculations accurately upon those for the surface of the Earth, assuming that the Earth is a sphere, and later correcting for the ellipticity of the great circle through the point on the central line and the Sun. The compression in this direction has to be reduced by a factor $1 - \cos^2 \phi \sin^2 A$ where ϕ is the latitude and A the azimuth of the Sun. For the northern and southern limits the accurate formulæ assume a far

* *Phil. Trans. Roy. Soc.*, A, 764, 255, 1937.

† *M.N.*, 92, 413, 1932.

too complicated form for easy calculation, as at any time the point on the northern limit of the path of totality on the Earth is not on the same generator of the shadow cone as the corresponding point of the northern limit of the path at a height above the surface.

The calculations have thus been made accurately by the simple expedient of expanding the radius of the Earth in the appropriate ratio, and substituting this increased value in the formulæ originally given by Chauvenet and expressed in a form suitable for calculating machines by Dr. L. J. Comrie

TABLE I

Latitude	Equatorial Radius		Correction to Latitude	
	100 km.	300 km.	100 km.	300 km.
± 0	1.01 568	1.04 703	± 0.000	± 0.000
10	.01 568	.04 704	.001	.003
20	.01 568	.04 705	.002	.006
30	.01 569	.04 707	.003	.008
40	.01 570	.04 710	.003	.009
± 50	1.01 571	1.04 713	± 0.003	± 0.009
60	.01 572	.04 715	.003	.008
70	.01 572	.04 717	.002	.006
80	.01 573	.04 719	.001	.003
90	.01 573	.04 719	.000	.000

in *M.N.*, 93, 175-181. They are consequently independent of the computed eclipse track for the surface of the Earth. In order to obtain a theoretical accuracy of 0°.001 in the calculations, which accuracy is almost certainly illusory with the present state of our knowledge of the positions and diameters of the Sun and Moon, it is necessary to increase the equatorial radius of the Earth by a ratio which depends on the latitude of the final position, so as to make the height at that latitude equal to the height required, and further to apply a correction to the final latitude to convert the latitude at a point on the ionosphere to the latitude of its corresponding projection on the Earth. These corrections are derived by simple considerations of the figure of the Earth, and are given by the following formulæ, from which Table I has been constructed.

If x , in units of the Earth's equatorial radius, is the height in latitude ϕ , and $S = \frac{\rho \sin \phi'}{\sin \phi}$, $C = \frac{\rho \cos \phi'}{\cos \phi}$ are the lengths of the normals to the Earth's surface at that point (tabulated in Table XII of the *Nautical Almanac* for 1939), the equatorial radius to be used is given by

$$1 + x \left(\frac{\sin^2 \phi}{S} + \frac{\cos^2 \phi}{C} \right)$$

and the correction, to diminish the computed latitude, by

$$\frac{x}{2} \left(\frac{1}{S} - \frac{1}{C} \right) \sin 2\phi.$$

South America.

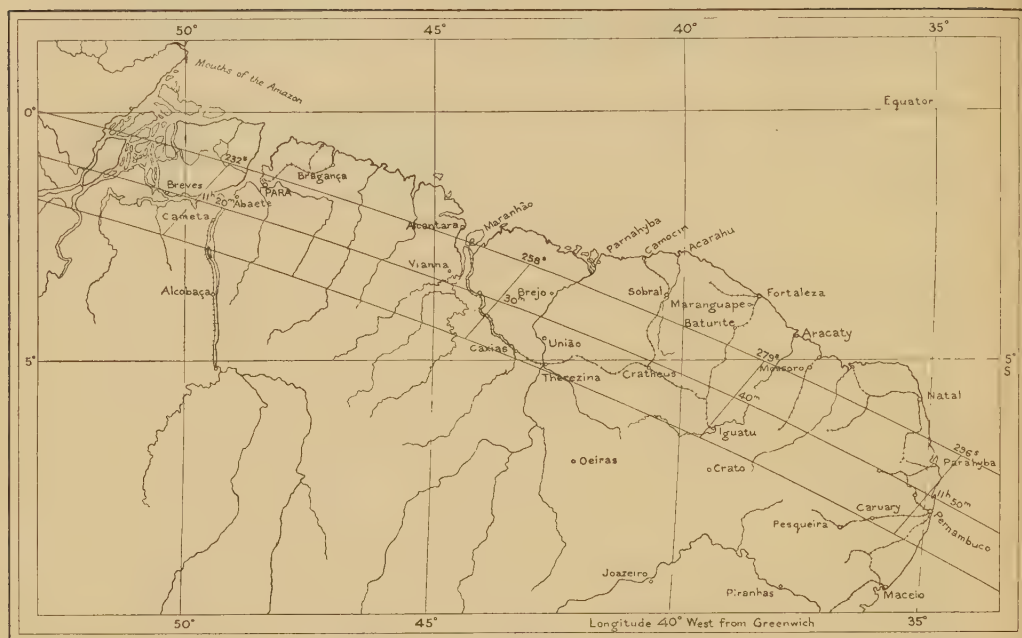


FIG. 1.—Track at a height of 100 kilometres.

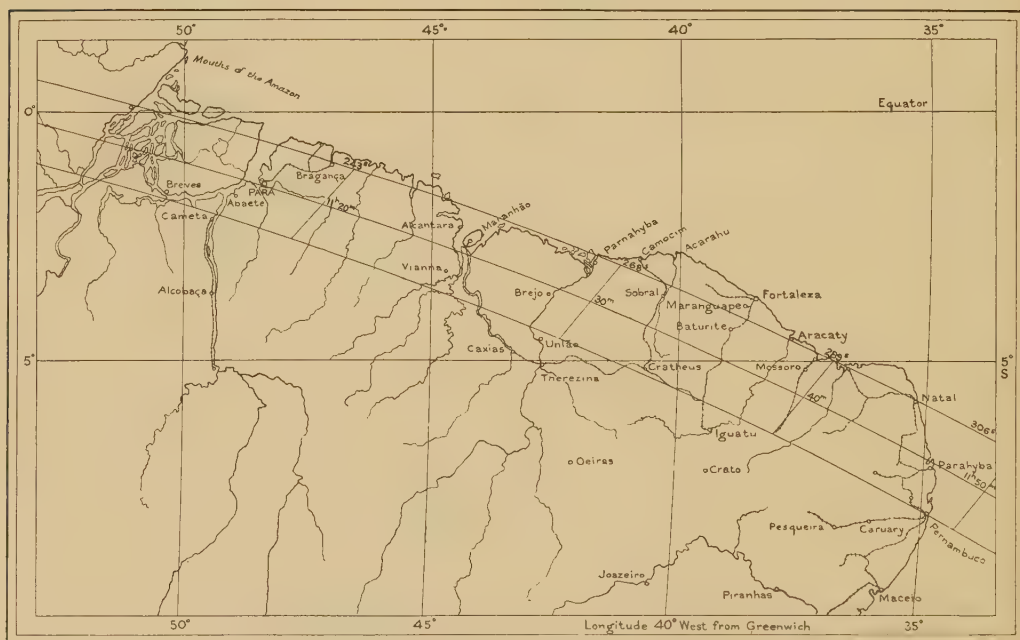


FIG. 2.—Track at a height of 300 kilometres.

South Africa.

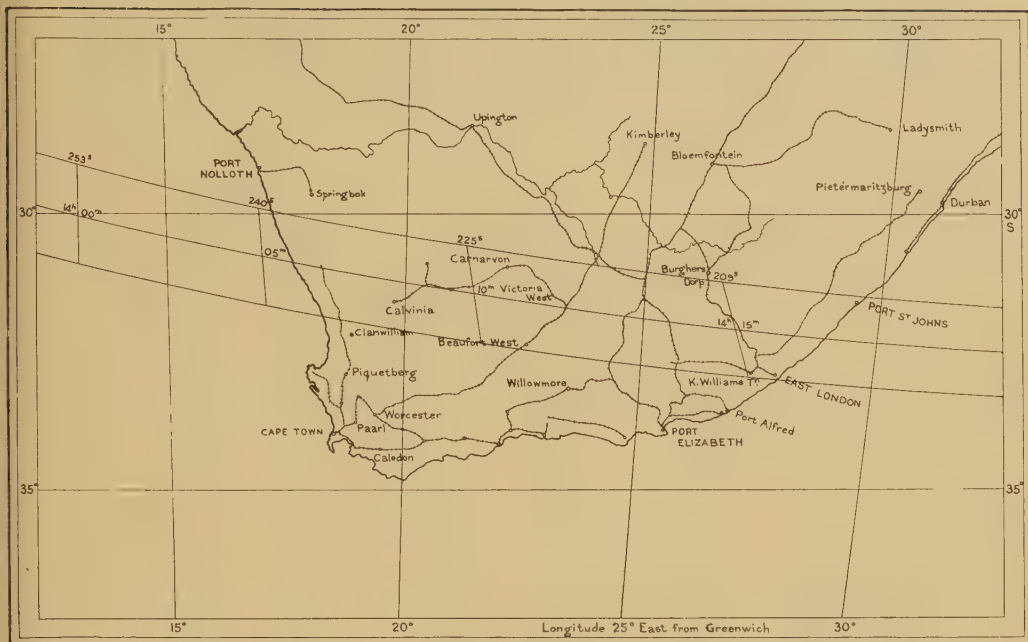


FIG. 3.—Track at a height of 100 kilometres.

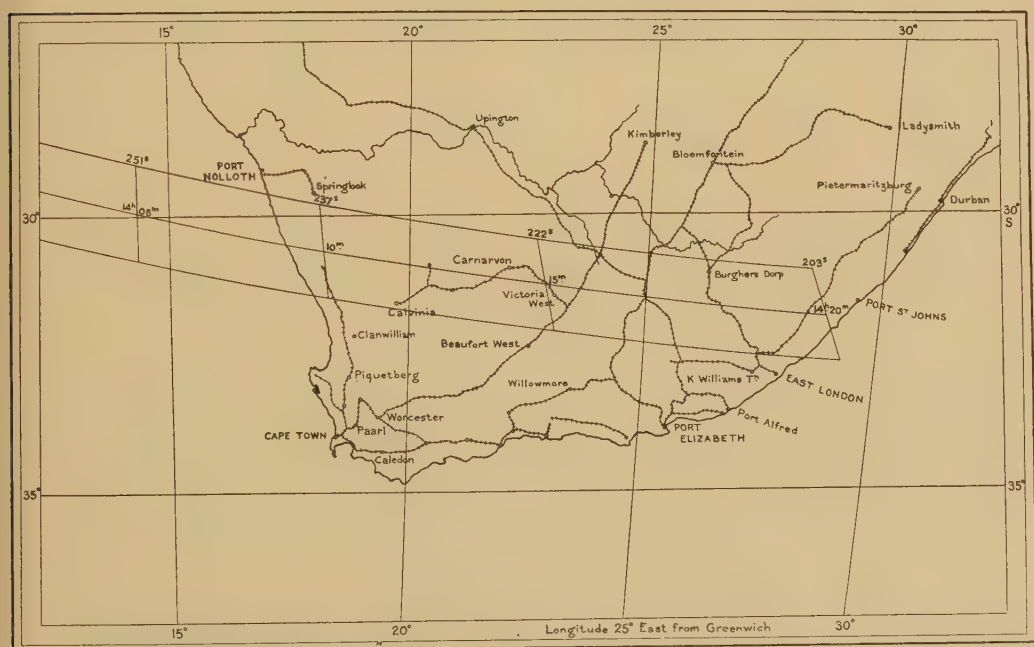


FIG. 4.—Track at a height of 300 kilometres.

The paths of the central lines have been checked by the differential computations from the track on the surface of the Earth, using the computed altitude and azimuth of the Sun, while those for the northern and southern limits have been checked by differencing and intercomparison. A further check has been applied on the general method by the rigorous deduction of the northern and southern limits from those on the Earth's surface, in

TABLE II

U.T.	Northern Limit		Central Line		Southern Limit		Duration	
	Latitude	Longitude	Latitude	Longitude	Latitude	Longitude		
At a height of 100 kilometres								
h m	° '	° '	° '	° '	° '	° '	m	
11 10	+ 1 12	+57 58	+ 0 31	+58 45	- 0 10	+59 35	3 19	
11 20	- 1 01	49 07	- 1 44	49 47	2 27	50 28	3 52	
11 30	3 03	43 06	3 48	43 43	4 32	44 21	4 18	
11 40	5 00	38 20	5 45	38 57	6 31	39 34	4 39	
11 50	6 53	34 20	7 40	34 56	8 26	35 32	4 56	
12 00	- 8 44	+30 47	- 9 31	+31 22	-10 19	+31 58	5 11	
...								
14 00	-29 08	-13 11	-30 01	-13 09	-30 55	-13 08	4 13	
14 05	29 53	16 53	30 46	16 56	31 39	17 01	4 00	
14 10	30 35	21 14	31 27	21 24	32 20	21 36	3 45	
14 15	31 14	26 37	32 05	26 57	32 56	27 19	3 29	
14 20	-31 47	-34 01	-32 35	-34 39	-33 24	-35 22	3 08	
At a height of 300 kilometres								
11 10	+ 0 56	+54 01	+ 0 15	+54 43	- 0 25	+55 27	3 33	
11 20	- 1 09	46 35	- 1 51	47 13	2 34	47 52	4 03	
11 30	3 06	41 12	3 50	41 48	4 33	42 26	4 28	
11 40	4 58	36 52	5 43	37 27	6 27	38 04	4 49	
11 50	6 47	33 09	7 33	33 44	8 18	34 20	5 06	
12 00	- 8 35	+29 51	- 9 21	+30 25	-10 07	+31 00	5 21	
...								
14 00	-28 21	-11 03	-29 12	-10 58	-30 05	-10 55	4 24	
14 05	29 04	14 20	29 56	14 20	30 48	14 21	4 11	
14 10	29 47	18 08	30 38	18 13	31 29	18 19	3 57	
14 15	30 26	22 40	31 16	22 52	32 07	23 06	3 42	
14 20	-31 02	-28 26	-31 51	-28 48	-32 40	-29 15	3 23	

the two cases of minimum altitude. Mainly as a matter of interest and to develop accurate methods of computation, the calculations have been carried to an accuracy of 0°.001 in each co-ordinate, though it is realised that for the purpose for which these are to be used, more approximate calculations would have sufficed. A correction of +1".5 has been applied to the mean longitude of the Sun as given by Newcomb's tables, while corrections of +2".0 and -0".5 have been applied respectively to the mean longitude and latitude of the Moon as given by Brown's tables.

Table II gives the geographical co-ordinates of the points on the central line, and northern and southern limits, at intervals of 10 minutes over South America and at intervals of 5 minutes over South Africa, to an accuracy of

1'. The portions of the tracks given cover only those land areas where observations are likely to be made. Accurate interpolation of these figures is not possible at low altitudes, but as it is these portions of the track which are least likely to be used, it was not considered necessary to calculate at a smaller interval. The column "duration" refers to the duration of totality on the central line; the altitude of the Sun has no great significance and is therefore not given.

The four maps show the tracks over South America and South Africa and have been drawn from the figures given in Table II. The information shown on the maps has been chosen to give a general idea of the locality covered by the tracks, and consists of the principal towns, rivers and railways and all the larger places within the belts of totality.

This work has been undertaken and the results published on the suggestions of Professor E. V. Appleton and the Astronomer Royal.

TOLERANCES PERMISSIBLE IN FLATS FOR AUTOCOLLIMATION TESTS.

C. R. Burch, B.A., Ph.D.

1. It is the purpose of this note to give formulæ for the amount of spherical convexity or concavity permissible in a flat used for autocollimation tests on lens- or mirror-systems in terms of the error permissible in the figure for which the test is null.

The tolerance depends on the "offence against the sine condition" of the system, the "coefficient of offence" being the quantity K in the expression

$$h = F \sin \theta [1 + K \sin^2 \theta + \dots \text{etc.}] \quad (1)$$

connecting the height h of a ray parallel to the axis: the focal length F and the angle θ at which the system converges that ray to the common focus. (Terms in higher powers of $\sin \theta$, in the bracket, may be neglected.) Let D be the diameter of the system, and δ the depth of (spherical) concavity of the flat in diameter D . Let the height of the centre and edge of the surface on autocollimation test be correct by definition, and denote the height excess, over true figure, of the zone of diameter $D/\sqrt{2}$ by ϵ , if the surface in question is a reflecting surface, and by $-2\epsilon/(\mu - 1)$ if it is a surface of index μ refracting into air.

Then a null test will be obtained when

$$\delta = 64 \left(\frac{F}{D} \right)^2 \frac{\epsilon}{(4K - \frac{1}{2})}. \quad (2)$$

Practical interest will be confined almost always to two cases—either that of a paraboloid, for which $K = +\frac{1}{4}$, or an aplanatic system, for which $K = 0$. Accordingly we can take the limits of tolerance for practical purposes to be

$$\delta = \pm 128 \left(\frac{F}{D} \right)^2 \epsilon. \quad (3)$$

Thus, even if we adopt the severe standard of $\epsilon \leq 2.5 \times 10^{-7}$ inch, we can allow a concavity in the flat of $\pm 3.4 \left(\frac{F}{D} \right)^2$ Newton rings (in green mercury light). So the permissible spherical concavity varies from ± 14 rings for a paraboloid of $F/2$ to the enormous amount of 750 rings for an objective of the not unusual ratio $F/15$.

2. *Measurement of K .*— K may be simply determined on the autocollimation bench, by knife-edge tests, as follows. Let the system be adjusted for autocollimation, the flat being squared-on to the optic axis,

and the final image being formed conveniently close to the source. Diaphragm the system so as to expose three zones only: a paraxial zone and two zones at opposite ends of a horizontal diameter D . Suppose the optic axis to be horizontal. Measure the longitudinal coma—that is, the amount by which the intersection of the ray from the zone to the observer's left with the paraxial ray lies nearer to him than the intersection of the ray from the zone on his right with the paraxial ray. Denote this distance by x_0 . (It is zero in a correctly figured, correctly squared-on system.) Now rotate the flat about a vertical axis, so that the image moves a distance ρ to the observer's left. It is permissible, but not essential, after measuring ρ to move the source until the source-image separation is again small. Measure the longitudinal coma again; let it now be x_ρ .

Then

$$x_\rho - x_0 = 3\rho \frac{D}{F} K. \quad (4)$$

3. *Derivation of Equations (2) and (4).*—If l' , l denote axial displacements in object and image spaces of refractive index unity, and θ' , θ conjugate angles of convergence, and if $l' \sin \theta'$ be infinitesimal, application of Clausius' theorem gives

$$l = \frac{l' \sin \theta' d\theta'}{\sin \theta d\theta} \quad (5)$$

(see, e.g., E. T. Whittaker, *The Theory of Optical Instruments* (1907, C.U. Press), p. 45).

If u' , u denote object and image distances from the appropriate principal planes, we shall have, when u' is sufficiently large,

$$\sin \theta' = \frac{u}{u'} \sin \theta [1 + K \sin^2 \theta + \text{etc.}]. \quad (6)$$

Identifying l , l' with du and $-du'$, and combining (5) and (6) we have

$$\frac{du}{u^2} = -\frac{du'}{u'^2} [1 + K \sin^2 \theta + \text{etc.}] [1 + 3K \sin^2 \theta + \text{etc.}] \cos \theta \sec \theta'. \quad (7)$$

Neglecting θ^4 compared with unity this can be written

$$\frac{du}{u^2} = -\frac{du'}{u'^2} \left[1 + \left\{ 4K - \frac{1}{2} \left(1 - \frac{u^2}{u'^2} \right) \right\} \theta^2 \right]. \quad (8)$$

We may further neglect u^2/u'^2 compared with unity, and regard the denominator of the L.H.S. as having the constant value F , in integrating from $u' = \infty$ to $u' = D^2/8\delta$ (the radius of curvature of the "flat"). This gives

$$\frac{\Delta F}{F^2} = -8\delta \left[1 + \left(4K - \frac{1}{2} \right) \theta^2 \right], \quad (9)$$

and from this expression for the longitudinal aberration of the image expression (2) follows.

Expression (4) is obtained by applying Clausius' equation to a small displacement perpendicular to the axis; the magnification of the system

in this direction is $\frac{\cos \theta' d\theta'}{\cos \theta d\theta} = \frac{u}{u'}(1 + 3K \sin^2 \theta \dots)$, and from this the

expression for the longitudinal coma on autocollimation follows simply.

Summary.—Limits are obtained for the tolerance of spherical convexity or concavity permissible in a flat used for autocollimation. Only the focal ratio of the system and its offence against the sine condition are involved; the determination of the "coefficient of offence" by knife-edge measurements is described.

OBSERVATIONS OF NOVÆ, 1937-1938.

W. H. Steavenson.

The observations recorded in this paper were, unless otherwise stated, made with the 20½-inch Hindle reflector. References to the comparison stars used will be found in earlier papers, and especially in *M.N.*, 92, 720, 1932.

Nova Aquilæ (1918)

Observed on six nights. This nova is perhaps very slightly variable. The apparent fall in magnitude, as compared with the estimates of recent years, is due to the adoption of Mitchell's magnitudes for the comparison stars (*Publ. Leander McCormick Obs.*, 6, 277), and these will always be used in future. There has been no real change in the mean brightness of the nova for many years.

Date	Mag.	Date	Mag.
1937 June 4	11.2	1937 July 18	11.0
20	11.1	Sept. 5	11.3
27	11.0	Oct. 5	11.1

Nova Aquilæ (1936, No. 1)

Observed on twenty-six nights. The observations indicate a slow fading, with little or no fluctuation.

Date	Mag.	Date	Mag.
1937 June 4	11.1	1937 July 13	11.3
7	11.1	14	11.4
8	11.1	16	11.4
13	11.1	18	11.5
15	11.0	Aug. 2	11.5
16	11.2	5	11.6
17	11.1	6	11.6
20	11.2	15	11.5
21	11.1	21	11.5
25	11.2	31	11.7
27	11.2	Sept. 5	11.5
29	11.2	Oct. 4	11.9
July 9	11.3	21	11.9

Nova Aquilæ (1936, No. 2)

Observed on nineteen nights. The nova is fading slowly, but with marked fluctuations, the range of which exceeds one magnitude. The magnitudes of the comparison stars are taken from a list issued by the British Astronomical Association.

Date	Mag.	Date	Mag.
1937 June 4	13.4	1937 July 14	13.7
7	13.7	16	13.5
8	13.4	18	13.4
13	13.5	Aug. 2	14.2
16	13.7	5	13.5
17	13.4	6	13.9
25	13.8	15	14.4
29	14.0	Sept. 3	13.7
July 9	13.3	5	13.8
13	13.6		

Supernova Can. Ven. (1937)

Observed on five nights. This object appeared in the nebula I.C. 4182. The estimates were made with the $5\frac{1}{2}$ -inch Cooke refractor and depend on two comparison stars, the magnitudes of which have been adopted as given in the Draper Catalogue. The stars are $+37^{\circ}2360$ ($8^m.3$) and $+38^{\circ}2398$ ($8^m.7$).

Date	Mag.
1937 Sept. 3	8.3
5	8.3
6	8.3
8	8.5
10	8.6

Nova (T) Coronæ (1866)

Observed on one night, 1937 June 5, when the magnitude was estimated to be 9.81. This is rather brighter than the estimates of recent years, and the slight increase in light indicated is probably real. The nova is distinctly yellowish in colour.

Nova Cygni (1876)

Observed on four nights. This nova has in past years exhibited a definite variability, with a total range of nearly a magnitude. If such variability still exists the few estimates obtained in 1937 do not happen to confirm the fact.

Date.	Mag.
1937 June 4	15.00
8	15.00
13	15.00
Sept. 5	14.91

Nova Cygni (1920)

Observed on one night, 1937 May 5, when the magnitude was estimated to be 15.4. The nova appears to have reached a stable minimum and there has been little or no change in its brightness for several years.

Nova Herculis (1934)

Observed on forty-one nights. The observations show a slow and practically steady fading. The nova has continued to exhibit a vivid green tint throughout the period. The magnitude estimates were made with the 3-inch and $5\frac{1}{2}$ -inch refractors.

Date	Mag.	Date	Mag.	Date	Mag.
1937 May 5	8.0	1937 June 17	8.2	1937 Oct. 3	8.4
10	8.0	20	8.3	4	8.3
21	8.0	21	8.1	21	8.3
24	8.1	24	8.0	1938 Mar. 18	8.5
27	8.2	25	8.1	19	8.5
31	8.2	29	8.1	31	8.5
June 4	8.1	Aug. 21	8.3	Apr. 1	8.6
5	8.2	Sept. 3	8.3	7	8.5
6	8.2	6	8.3	10	8.5
7	8.1	10	8.4	11	8.5
8	8.2	15	8.3	12	8.5
13	8.2	20	8.2	15	8.5
15	8.1	23	8.3	17	8.5
16	8.2	Oct. 1	8.3		

Nova Lacertæ (1910)

Observed on two nights (1937 June 4 and September 5), on both of which the magnitude was estimated to be 14.4. This nova appears to have reached a stable minimum, and no definite changes in its light have been observed for many years.

Nova Lacertæ (1936)

Observed on fifty-one nights. The observations show a sensibly steady fading. The magnitudes adopted for the comparison stars were those issued by the British Astronomical Association. On 1937 August 15 the nova appeared like a small planetary nebula, about one second of arc in diameter, and was greenish in colour.

Date	Mag.	Date	Mag.	Date	Mag.
1937 May 24	10.8	1937 July 9	11.2	1937 Sept. 8	11.5
27	10.8	13	11.3	10	11.5
June 4	10.8	14	11.2	20	11.7
5	10.9	16	11.2	26	11.8
6	10.9	18	11.3	Oct. 1	11.7
7	10.9	Aug. 2	11.3	4	11.7
8	11.0	3	11.4	5	11.8
13	10.9	5	11.3	9	11.7
15	11.0	6	11.2	21	11.8
16	10.9	15	11.4	30	11.9
17	10.9	21	11.5	Nov. 9	12.1
20	11.2	22	11.4	14	12.0
21	11.0	25	11.4	Dec. 5	12.5
24	11.0	28	11.5	12	12.3
25	11.2	31	11.6	1938 Jan. 9	12.5
27	11.1	Sept. 5	11.5	Feb. 1	12.8
29	11.1	6	11.5	Mar. 3	12.8

Nova Ophiuchi (1848)

Observed on ten nights. This nova is slightly variable, but the few observations obtained in 1937 do not bring out so large a range as has been observed in previous years.

Date	Mag.	Date	Mag.
1937 June 4	12.7	1937 June 13	12.7
5	12.6	29	12.8
6	12.6	July 9	12.8
7	12.6	13	12.8
8	12.5	14	12.8

Nova Persei (1901)

Observed on forty nights. The normal variability of this nova, which had been much modified in the period 1936-1937, has been resumed, the observations showing a total range of 1^m.56.

Date	Mag.	Date	Mag.	Date	Mag.
1937 Aug. 7	12.89	1937 Oct. 5	13.02	1938 Mar. 1	12.95
15	12.89	8	13.57	2	12.95
Sept. 3	13.10	17	13.77	3	12.95
5	13.02	25	13.52	4	12.99
6	12.94	30	13.05	5	13.05
10	12.79	Nov. 1	12.84	12	12.69
15	13.42	9	13.00	18	12.53
20	13.62	11	13.36	20	12.69
26	13.57	14	13.57	29	12.32
28	13.57	25	13.88	31	12.48
29	13.72	29	13.82	Apr. 1	12.64
Oct. 1	13.21	1938 Jan. 25	13.87	17	12.59
3	13.16	Feb. 1	13.87		
4	13.10	3	13.77		

Supernova Persei (1937)

Observed on eight nights. This nova appeared near the following end of the nebula N.G.C. 1003. The comparison stars used were among those given by Beer and Gregory in *M.N.*, 98, 217, 1938. The following are my estimates of the visual magnitudes of those which I used. They are probably too bright by about half a magnitude. The letters are those of Beer and Gregory (*loc. cit.*).

$$h = 12.5, \quad j = 12.3, \quad l = 12.8, \quad n = 13.8, \quad w = 14.5.$$

Date	Mag.	Date	Mag.
1937 Sept. 15	12.0	1937 Sept. 29	12.5
20	12.2	Oct. 1	12.8
23	12.0	3	12.9
26	12.3	4	13.0

ON THE DISTRIBUTION OF ABSOLUTE MAGNITUDES IN THE VICINITY OF THE SUN.

W. J. Luyten.

1. The proper motion survey of the southern hemisphere undertaken by the writer in co-operation with the Harvard Observatory has now been completed. The material consists of pairs of plates taken with the Bruce 24-inch refractor, with a scale of 1 mm. = 1'. Altogether 1009 plates are required to cover the southern hemisphere completely; 950 old plates were available, while the remainder was not taken owing to the cloudy season at Arequipa. These missing regions lie largely between 0° and -10° declination and 3–8 hours R.A., an area which is accessible from observatories in the northern hemisphere. All new plates have been taken, 300 by the writer, the remainder by the Harvard Observatory; they have all been blinked by the writer. In general the plates show stars of the sixteenth photographic magnitude, but about 100 plates of long exposure go down to 18. The average interval is about thirty years; motions as small as $0''.050$ annually are easily detected in the central part of the plates. An area of about 30 square degrees was blinked on each plate; owing to overlap all but 1232 square degrees—or only 6 per cent. of the southern hemisphere—has been covered. In this area a total of 94,300 stars with appreciable proper motion have been found.

2. One of the principal reasons for undertaking the proper motion survey was a desire to obtain a new derivation of the distribution of absolute magnitudes in space. Knowledge of the numbers of stars with large proper motion is essential for this.

(a) Altogether 855 stars were found with proper motions in excess of $0''.5$ annually, counting double stars as one; among these 683 stars are brighter than the magnitude 14.5 photographic, and 172 stars are fainter than this limit.

(b) Comparisons with other determinations—though admittedly scanty—indicate only a negligible correction to the magnitudes estimated by the writer.

(c) Making allowance for the unblinked areas and for stars overlooked would bring the total number of stars with $\mu > 0''.5$ per annum and with m brighter than 14.5 to 710 for the entire southern hemisphere.

(d) Ninety-two of the long-exposure plates were judged to be of first-class quality. From the number of proper motion stars found in the central twenty square degrees of these plates—1840 square degrees or 9 per cent. of the southern hemisphere in all—the total number of stars between 14.5 and 16.5 and with $\mu > 0''.5$ per annum is estimated to be 603 for the southern hemisphere. There is, however, considerable uncertainty

attached to this number, since the long-exposure plates are far from uniformly distributed over the sky. Below the limit of 16.5 ptg. the survey is woefully incomplete, but the indications are that the number of stars with large proper motion falls off very sharply below 17^m.

3. Comparison with data in the northern hemisphere which are known to be incomplete below 11^m ptg. indicate some puzzling differences. The brightest stars—brighter than 7.5 ptg.—appear to have *larger* motions in the south than in the north, while stars with very large motions and fainter than 9.5 ptg. seem to be more numerous in the north. Owing to the lack of homogeneity in the northern survey no definite conclusions can be reached, but the indications are that the differences are larger than the permissible statistical fluctuations.

4. In the southern hemisphere there is, of course, a conspicuous dearth of stars with $\mu > 0''.5$ in the area surrounding the solar antapex, near 7^h, -30°. Almost equally conspicuous is the absence of these large proper motion stars in the region around 19^h, -30°, not far from the vertex.

5. Before an attempt can be made at deriving the luminosity function from these data, two further kinds of information are required, viz.

(a) The mean parallax of stars of given m and μ , and

(b) The dispersion of the true parallaxes around this mean.

Using the trigonometric parallaxes of 620 stars with motions in excess of 0''.5 annually, we find a nearly linear relation between the absolute magnitude m and the reduced proper motion $H = m + 5 + 5 \log \mu$, with a dispersion in M of 1.22. Another, and perhaps more accurate, representation of the regression of M on H is given by an empirical formula with a dispersion in M increasing from 1.0 for the brightest stars to 1.35 for the faintest.

Applying these relations to the stars with $\mu > 0''.5$ we calculate, for each subdivision of m and μ , the fraction which have parallaxes exceeding 0''.20, and in this way derive the luminosity function for stars nearer than 5 parsecs.

While this is the customary procedure it may be emphasized that its results are subject to large uncertainties. What it amounts to, in practice, is the calculation of the fraction of these stars of large angular motion which have linear tangential velocities less than a given amount. Thus, for stars, say, between 12.5 and 13.5 and $\mu = 0''.5$ only those which have tangential velocities less than 12 km./sec. contribute to the luminosity function of stars nearer than five parsecs. The usual procedure is to determine the mean error, *i.e.* the root-mean-square deviation $M - \bar{M}$, or, what amounts to the same thing, that in $\log p/\bar{p}$, and assume that the actual distribution follows a normal error curve. Such data as we have at our disposal do indeed support this assumption, but it should be borne in mind that the principal contributions to the luminosity function are made by those stars whose parallaxes exceed the mean of their group by five or more times, *i.e.* we are dealing with the very small frequencies in the "tail" of a normal error curve at a distance from the mean three times the dispersion. It is evident that even small uncertainties in the original data will become greatly magnified in the final frequencies of the luminosity function. It might be

thought that the frequencies of stars of small tangential velocities might be better calculated directly from the velocity distribution, but in practice this will amount to the same thing since no radial velocities are known for these faint stars, and our knowledge of the distribution of their velocities must be derived from the same parallaxes employed here.

6. By slightly varying the constants used in the regression formula for M on H several different luminosity functions were derived. Fortunately, it is seen that the most serious uncertainty—that of the functional form of the distribution $\log p/\bar{p}$ —affects chiefly the star density and much less the form of the luminosity law. The curve finally adopted strikes an average between the several trials, and represents, I believe, about the best we can expect at present. In Table I are given the actual numbers (smoothed,

TABLE I

$M.$	$N.$	$M.$	$N.$	$M.$	$N.$
- 1	0.001	8	1.24	17	1.95
0	0.023	9	1.49	18	1.21
1	0.11	10	1.81	19	0.5 :
2	0.20	11	2.11	20	0.1 :
3	0.29	12	2.38	21	0.02 ::
4	0.40	13	2.65		
5	0.57	14	2.84		27
6	0.76	15	2.84		
7	0.97	16	2.51		

Numbers given are actual numbers of stars nearer than five parsecs and in the southern hemisphere. Magnitudes are photographic.

of course) of stars of different absolute magnitudes nearer than five parsecs in the southern hemisphere. The total number in the sky should thus be 54, corresponding to a star density of 0.1 star per cubic parsec.

A graph of the luminosity curve shows that the curve reaches a maximum around $M = +14.7$ ptg., and is distinctly skew as it has a much slower rise than fall. Partly this may be due to the fact that the material from which the present curve is derived is known to be incomplete below apparent magnitude 16.5; hence all frequencies below $M = 18$ are likewise too small. It may well be that the deficiency of the present curve is already appreciable for stars brighter than $M = 18$, since, *e.g.*, the curve indicates that there should be only four stars fainter than $M = 16.5$ ptg. nearer than five parsecs, while two are now known, and several more may exist.

7. Using the best available data on the relation between mass and luminosity, and making due allowance for the occurrence of double stars and white dwarfs, we find that the present luminosity curve indicates a density of 0.060 Sun per cubic parsec. Watson has estimated the density of meteoric dust in space to be of the order of $0.015 \odot .ps.^{-3}$, thus giving a total of $0.075 \odot .ps.^{-3}$ in the vicinity of the Sun. While it is difficult to calculate how much mass, if any, is represented by bodies intermediate between the faintest star considered here—of photographic absolute magnitude +20—and the largest meteor, we may perhaps estimate the

total mass to be of the order of $0.08 \odot . \text{ps.}^{-3}$. From gravitational considerations Oort finds this total mass to be $0.089 \odot . \text{ps.}^{-3}$. It appears tempting to conclude, therefore, that the present curve is reasonably complete down to, at least, $M = +16$.

A more detailed calculation and discussion will be published shortly in the Publications of the Astronomical Observatory of the University of Minnesota ; only a summary is given here which it is hoped may be of some use to statistical workers.

In concluding, it must be emphasized again that the parallax material is still far too incomplete to warrant anything but provisional conclusions, and it will be many years before we can speak with anything approaching finality about the true form of the luminosity function even in the immediate vicinity of the Sun.

Summary.—The stars with proper motions larger than $0''.5$ annually which were obtained from the Bruce Proper Motion Survey in the Southern Hemisphere have been used to derive the luminosity function. The formula for mean parallax and the dispersion of the true parallaxes around the mean have been obtained from trigonometric parallax material. The resulting luminosity curve shows a maximum around photographic absolute magnitude $+15$ and appears to possess a steeper descent than rise. The total mass in the vicinity of the Sun is estimated at 0.08 Sun per cubic parsec.

Leiden :
1938 April 6.

Erratum

M.N., 98, 288 (Helwan Observatory Report), range in temperature of clock room, for $25^{\circ}.4$ C., read $19^{\circ}.6$ C.